Single-Molecule FRET Spectroscopy of Intrinsically Disordered Proteins in Crowded Solutions and in Living Mammalian Cells

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Iwo König

von

Madiswil BE

Promotionskommission

Prof. Dr. Benjamin Schuler (Vorsitz)

Prof. Dr. Andreas Plückthun

Prof. Dr. Achillefs Kapanidis

Zürich, 2017

Abstract

Single-molecule spectroscopy is a powerful tool for quantifying the conformational heterogeneity and structural dynamics of both folded and intrinsically disordered proteins (IDPs). IDPs lack ordered structure under native conditions and are typically characterized by low hydrophobicity and high net charge. A high fraction of eukaryotic proteins either contain disordered regions or are completely unstructured. IDPs are often part of dynamic intracellular networks of interacting proteins in signaling and regulatory processes. Due to their flexibility, IDPs adopt expanded conformations which are expected to be more sensitive to effects of the intracellular environment, such as macromolecular crowding, than folded proteins. Crowding is a steric exclusion effect of macromolecules on other molecules and can affect dimensions, conformational stabilities and folding kinetics of proteins. We have thus studied the effect of co-solutes on the dimensions and dynamics of IDPs *in vitro* and in living mammalian cells using single-molecule Förster resonance energy transfer (FRET) spectroscopy.

First, *in vitro* experiments were performed to study the effect of artificial crowders such as polyethylene glycol (PEG) on different IDPs. We observed a compaction of all IDPs not only with increasing PEG concentration, but also with increasing size of the crowders. This finding deviated even qualitatively from the predictions of scaled particle theory (SPT), the standard theory used to describe macromolecular crowding. However, we were able to rationalize the observed behavior quantitatively if we accounted for the polymeric nature of both the IDPs and the crowder in the theoretical description.

Next, we studied IDPs in living cells with single-molecule FRET spectroscopy. The application of single-molecule FRET spectroscopy in living mammalian cells had remained challenging due to several shortcomings, such as the design and reproducible delivery of labeled molecules, the range of applicable analysis methods, and suboptimal cell culture conditions. We addressed these limitations by combining confocal single-molecule FRET spectroscopy with microinjection and by choosing fluorescent dyes which are shifted towards larger emission wavelengths. With this integrated approach, we demonstrate the feasibility of probing IDP dimensions and dynamics down to the nanosecond regime in live eukaryotic cells. We illustrated the versatility of the approach by determining the dimensions and chain reconfiguration dynamics of the intrinsically disordered protein prothymosin α (ProTα); by analyzing the temperature dependence of the structural stability of frataxin; and by quantifying the folding dynamics of the small protein GB1. We found that the influence of the intracellular environment on the observed proteins was negligible. ProTα remained unstructured, and its dimensions were unaffected by the cell. Only the translational diffusion and the chain reconfiguration dynamics of the IDP were slowed down because of the increased intracellular viscosity. Also, the structural stability of frataxin and the folding kinetics of GB1 were not altered compared to *in vitro* experiments at physiological salt concentrations.

As a last step, we used in-cell single-molecule FRET spectroscopy methods to study intracellular crowding in further detail. Since the previously observed effects on IDP dimensions were surprisingly small, we increased crowding in HeLa cells by subjecting them to an osmotic shock to reduce the cell volume and thus increase the intracellular concentration of molecules. Indeed, ProTα was more compact in such

crowded cells than in cells at physiological conditions but its reconfiguration dynamics were unchanged. This effect was mimicked with PEG crowders and allowed us to estimate the equivalent macromolecular concentrations in the cytosol of HeLa cells under physiological conditions to be approximately 5 % (w/v), significantly lower than commonly assumed.

Single-molecule FRET spectroscopy thus allowed us to study structural and dynamic properties of IDPs and other proteins in living cells and under crowding conditions in vitro with high sensitivity and reproducibility. Even though IDPs were shown to be very sensitive to crowding in vitro, we found that the effect of the intracellular environment on the dimensions and dynamics of IDPs was small in HeLa cells, probably because these cells are less crowded than commonly assumed. With this newly established in-cell method, a broad range of questions can be addressed to analyze protein conformation, dynamics, and functionality in live eukaryotic cells.

Zusammenfassung

Einzelmolekülspektroskopie ist eine leistungsfähige Methode, um die strukturelle Heterogenität und Dynamik von gefalteten und intrinsisch unstrukturierten Proteinen („intrinsically disordered proteins“, IDPs) zu quantifizieren. IDPs besitzen unter nativen Bedingungen keine geordnete Struktur und sind typischerweise durch eine geringe Hydrophobizität und eine hohe Nettoladung charakterisiert. Ein grosser Teil der eukaryotischen Proteine enthält entweder grosse unstrukturierte Bereiche oder ist komplett unstrukturiert. IDPs sind oftmals Teil eines dynamischen intrazellulären Netzwerkes verschiedener interagierender Proteine, besonders in der Signalübertragung oder bei zellregulatorischen Prozessen. Aufgrund ihrer Flexibilität besitzen IDPs oft eine sehr expandierte Struktur, wodurch sie auf Einflüsse der intrazellulären Umgebung, wie z.B. makromolekulares Crowding, empfindlicher reagieren als gefaltete Proteine. Crowding ist ein sterischer Ausschlusseffekt, welcher durch Makromoleküle ausgeübt wird und die Dimensionen, die konformationelle Stabilität sowie die Faltungskinetik von Proteinen beeinflussen kann. Wir haben deshalb den Einfluss verschiedener Makromoleküle auf die Dimensionen und Dynamik von IDPs unter *in vitro*-Bedingungen und in lebenden Säugerzellen mittels Einzelmolekül-Förster-Resonanz-Energietransfer (FRET)-Spektroskopie untersucht.

Zuerst wurden in vitro-Experimente durchgeführt, um den Einfluss von künstlichen Crowdern, z.B. Polyethylenglykol (PEG), auf verschiedene IDPs zu untersuchen. Wir beobachteten eine Kompaktierung aller IDPs mit steigenden PEG-Konzentrationen und mit steigender Crowderlänge. Dieser Befund weicht von Vorhersagen der *scaled particle*-Theorie (SPT) ab, welche die Standardtheorie zur Beschreibung von makromolekularem Crowding ist. Wir waren jedoch in der Lage, das beobachtete Verhalten zu beschreiben, indem wir die polymeren Eigenschaften der IDPs und der Crowder in die theoretische Beschreibung miteinbezogen.

In einem weiteren Schritt untersuchten wir IDPs in lebenden Säugerzellen mittels Einzelmolekül-FRET-Spektroskopie. Dies stellt sehr hohe experimentelle Anforderungen, z.B. an das Probendesign und die Art und Weise, wie die fluoreszenzmarkierten Moleküle in die Zelle eingebracht werden, welche Analysemethoden in der Zelle angewandt werden können, sowie an die Zellkulturbedingungen. Wir lösten diese Probleme, indem wir konfokale Einzelmolekül-FRET-Spektroskopie zusammen mit Mikroinjektion benutzten, und indem wir fluoreszierende Farbstoffe mit längeren Emissionswellenlängen verwendeten. Mit diesem integrierten Vorgehen konnten wir zeigen, dass die Dimensionen von IDPs sowie deren Dynamik bis ins Nanosekunden-Regime in lebenden eukaryotischen Zellen untersucht werden können. Die Vielseitigkeit dieses Ansatzes wurde gezeigt, indem wir die Dimensionen und Rekonfigurationsdynamik des intrinsisch unstrukturierten Proteins Prothymosin α (ProTα) gemessen haben, indem wir die Temperaturabhängigkeit der Konformationsstabilität von Frataxin untersuchten, und indem wir die Faltungsdynamiken des kleinen Proteins GB1 quantifizierten. Anhand dieser Experimente wurde deutlich, dass der Einfluss des intrazellulären Milieus auf die untersuchten Proteine vernachlässigbar war. ProTα blieb unstrukturiert, und seine Dimensionen wurden in der Zelle nicht verändert. Einzig die Translationsdiffusion und die Rekonfigurationsdynamik des IDPs wurden aufgrund der erhöhten intrazellulären Viskosität geringfügig verlangsamt. Auch die

strukturelle Stabilität von Frataxin und die Faltungskinetik von GB1 verhielten sich nicht anders als in *in vitro*-Experimenten unter physiologischen Salzkonzentrationen.

In einem letzten Schritt verwendeten wir Einzelmolekül-FRET-Spektroskopie, um die Relevanz von intrazellulärem Crowding weiter zu untersuchen. Da der bis dahin beobachtete Einfluss der Zelle auf die Dimensionen von IDPs erstaunlich klein war, erhöhten wir das Crowding in der Zelle mittels osmotischem Schock. Der osmotische Schock reduziert das Zellvolumen, worauf die intrazelluläre Konzentration an Molekülen ansteigt. ProTα war in den so behandelten Zellen tatsächlich kompakter als in Zellen unter physiologischen Bedingungen, jedoch zeigte die Rekonfigurationsdynamik keine Änderung. Mit PEG als Crowder konnten wir ähnliche Bedingungen wie in der Zelle *in vitro* erzeugen, woraus wir schlussfolgern konnten, dass die äquivalente makromolekulare Konzentration im Zytosol von HeLa-Zellen unter physiologischen Bedingungen bei etwa 5 % (w/v) liegt, deutlich niedriger als üblicherweise angenommen.

Einzelmolekül-FRET-Spektroskopie erlaubte uns somit, mit hoher Sensitivität und Reproduzierbarkeit die strukturellen und dynamischen Eigenschaften von IDPs und anderen Proteinen in lebenden Zellen und unter Crowding-Bedingungen zu untersuchen. Obwohl IDPs *in vitro* sehr empfindlich gegenüber Crowding waren, beobachteten wir, dass der Einfluss des intrazellulären Milieus von HeLa-Zellen auf die Dimensionen und Dynamik von IDPs und anderen Proteinen überraschend klein war. Der Grund dafür ist offenbar, dass die Konzentration an Makromolekülen in diesen Zellen geringer ist, als gemeinhin angenommen wird. Mit der hier neu entwickelten Methode können zukünftig viele verschiedene Fragen in Bezug auf die Konformation, Dynamik und Funktion von Proteinen in lebenden eukaryotischen Zellen untersucht werden.

Publications

Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells
I. König, A. Zarrine-Afsar, M. Aznauryan, A. Soranno, B. Wunderlich, F. Dingfelder, J. C. Stüber, A. Plückthun, D. Nettels, B. Schuler
Nat Methods, 12(8), 773–779, 2015

Gas-phase FRET efficiency measurements to probe the conformation of mass-selected proteins
M. F. Czar, F. Zosel, I. König, D. Nettels, B. Wunderlich, B. Schuler, A. Zarrine-Afsar, R. A. Jockusch
Analytical Chemistry, 87(15), 7559–7565, 2015

The effects of stabilizing and directional selection on phenotypic and genotypic variation in a population of RNA enzymes
E. J. Hayden, S. Bratulic, I. König, E. Ferrada, A. Wagner
J Mol Evol, 78(2), 101–108, 2014

Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments
A. Soranno, I. König, M. B. Borgia, H. Hofmann, F. Zosel, D. Nettels, B. Schuler
Proc Natl Acad Sci USA, 111(13), 4874–4879, 2014

Reduction of carry over in liquid-handling systems with a decontamination step integrated in the washing procedure
M. Iten, R. Weibel, I. König, R. Beckbissinger, T. Benthien, W. Hälg, N. Ingenhoven, A. Lehnert, L. Oeltjen, C. Zaborosch
J Lab Autom, 15(5), 379–389, 2010

Development of standard test procedures for quantifying carry over from fixed pipetting tips in liquid-handling systems
R. Weibel, M. Iten, I. König, R. Beckbissinger, T. Benthien, W. Hälg, N. Ingenhoven, A. Lehnert, L. Oeltjen, C. Zaborosch
J Lab Autom, 15(5), 369–378, 2010

Respiratory syncytial virus subunit vaccine based on a recombinant fusion protein expressed transiently in mammalian cells
S. Nallet, M. Amacker, N. Westerfeld, L. Baldi, I. König, D. L. Hacker, C. Zaborosch, R. Zurbriggen, F. M. Wurm
Vaccine, 27(46), 6415–6419, 2009

Purification and structural characterization of 3-hydroxypropionaldehyde and its derivatives
S. Vollenweider, G. Grassi, I. König, Z. Puhan
J Agric Food Chem, 51(11), 3287–3293, 2003

Table of contents

1 Introduction

1.1. Macromolecular crowding *in vitro* and *in vivo*

1.1.1. What is macromolecular crowding?

Proteins, nucleic acids, polysaccharides or other macromolecules interact with other molecules in an attractive or repulsive manner. Attractive interactions, such as charge or hydrophobic interactions, can drive protein folding, lead to the formation of protein complexes or enables enzymatic reactions in biological systems. On the other hand, repulsive interactions are, per definition, inhibiting the interaction of macromolecules, but can still influence their dynamic or conformational behavior. Steric exclusion is one example for non-specific repulsive interactions, where volume gets occupied by macromolecules or other obstacles.

Macromolecular crowding is defined as a steric exclusion effect exerted by macromolecules in solution. The term *crowding* for steric exclusion effects was introduced by Minton and Wilf in 1981 (1), when they observed that the addition of high concentrations of an unrelated globular protein to a solution containing glyceraldehyde-3-phosphate dehydrogenase (GAPD) led to a preferred formation of tetrameric GAPD due to space-filling properties of the added species. The occupied volume, which depends on the concentration, the size and the shape of the crowder, is not available for other macromolecules and can thus influence them in many ways. Crowding can affect processes such as the free diffusion of macromolecules, association kinetics or the compactness of proteins (2) (Fig. 1).

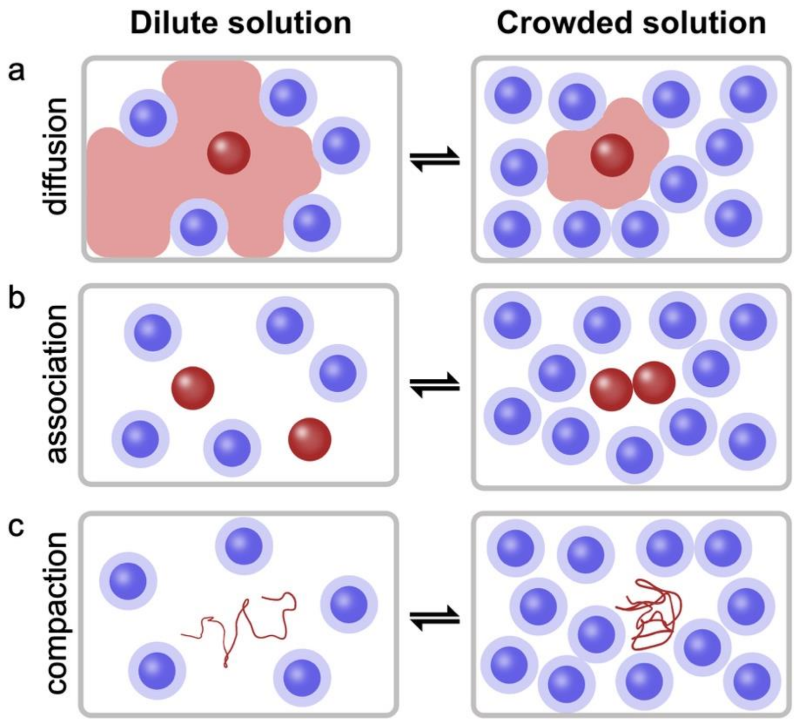


Fig. 1: Scheme of different effects of macromolecular crowding. (a) Macromolecular crowders (blue) occupy space and thus affect other macromolecules (red) in their diffusional behavior; (b) association kinetics; (c) or the compactness of proteins or other polymers.

1.1.2. Effects of crowding on proteins

It is generally thought that steric exclusion increases the stability of proteins because crowding favors more compact, folded states over extended unfolded states (3), and therefore also enhance association reactions (Fig. 1). Early experiments showed that, e.g., DNA ligation was stimulated (4) or that the solubility of deoxygenated sickle hemoglobin was reduced (5) in the presence of high concentrations of unrelated molecules. More theoretical and experimental studies have been conducted in the following decades on the influence of the high volume-occupancy of crowded solutions on proteins. A few studies indeed found that high concentrations of Ficoll, a branched polysaccharide, can slightly stabilize proteins. The presence of this crowder increased the melting temperature of RNase A and lysozyme (6), or of folded yeast phosphoglycerate kinase (PGK) (7) by a few Kelvin. A similarly small effect on the melting temperature was found when protein L was crowded with dextran, a linear polysaccharide (8). Those findings were in agreement with simulations, were only a modest stabilization of the binding of two protein subunits was found (9).

Along these lines, the influence of crowding on protein association rates was studied. Simulations of the binding of two subunits in the *E. coli* polymerase III holoenzyme showed a modest stabilization of the complex upon addition of dextran of Ficoll (9). Other experimental studies showed that crowding accelerates the self-association of human apolipoprotein C-II (10) and β-synuclein (11). The formation of amyloid-like fibrils by these proteins is relevant for several human diseases, such as Parkinson’s or Alzheimer’s disease.

Steric exclusion can induce structural changes in proteins. Crowding increased the enzymatic activity of phosphoglycerate kinase (PGK) by rearranging the subunits of the protein (7). Crowding can also lead to a compaction of proteins. The partially unfolded CORE domain of *E. coli* adenylate kinase was found to be more compact in the presence of high concentrations of dextran (12). A similar compacting effect of polyethylene glycol (PEG) was found on the unfolded states of ribonuclease A (13).

Increasing concentrations of crowders increase the viscosity of a medium and thus influence dynamic processes of proteins. It was shown that the increased viscosity, induced by glycerol, changed the rotational and translational diffusion of β-lactamase inhibitor protein, following a Stokes-Einstein dependence for diffusion as a function of viscosity (14). This relation did not hold anymore when a larger crowder (PEG 8000) was used. Also the association of β-lactamase and β-lactamase inhibitor protein in a PEG 8000 solution depended weakly on the viscosity of the solution, pointing out that the size ratio between proteins and crowders is relevant for protein diffusion and thus association reactions (15, 16).

Crowding experiments are usually carried out using artificial crowders such as dextrans (linear polysaccharide), Ficoll (branched polysaccharide) or PEG (linear polymer). A difficulty in analyzing crowding experiments is to exclude attractive interactions between the crowder and the protein. It was observed, e.g., that PEG interacted with hydrophobic residues of cytochrome c (17). Exposing the protein to the monomer of the artificial crowder is thus an important control for possible interactions. These competing attractive interactions can become even more dominant if proteins are chosen as crowders. BSA or lysozyme, e.g., were found to destabilize the protein chymotrypsin inhibitor 2 (CI2) (18) due to electrostatic or hydrophobic interactions.

1.1.3. Effect of the intracellular environment on proteins

Compared to the chemically defined conditions in a test tube, the inside of a cell is a complex ensemble of organelles, macromolecules, small metabolites and ions. Differential interference contrast (DIC) images of HeLa cells (Fig. 2a) show substructures within the cell, but identifying anything but large objects such as the nucleus is difficult. Using specific fluorescent probes, the staining of organelles or filaments provides a better insight into how a cell is organized (Fig. 2b). The complexity of the cytosol on the macromolecular level is illustrated in Fig. 2c, where the cytosol of an *E. coli* cell was simulated assuming a defined number of proteins and nucleic acids (19, 20). It is obvious that the cytosol or nucleus of cells is a highly crowded environment.

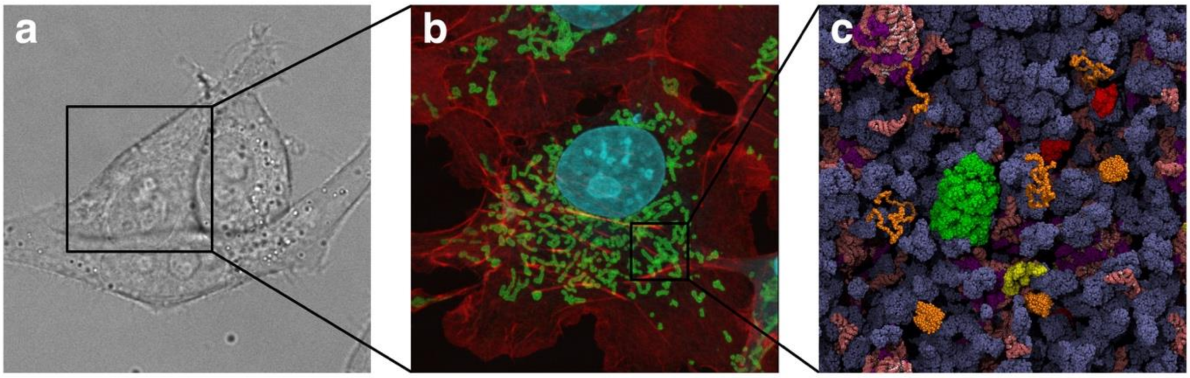


Fig. 2: Illustration of the complex organization of cells. (a) A DIC image of HeLa cells showing the coarse organization of organelles within a cell. (b) Using fluorescent probes, organelles are stained specifically, showing mitochondria (green), actin filaments (red) or the nucleus (cyan) of HeLa cells. (c) Illustrates of a simulated cytosol consisting of different proteins and nucleic acids (19).

Estimating protein or nucleic acid concentrations within cells is difficult. The macromolecular concentrations in *E. coli* were estimated to be 300 – 400 mg/ml (21). Estimates of cytosolic macromolecular concentrations in eukaryotic cells range from 50 – 200 mg/ml (22, 23). Assuming macromolecular concentrations within this range, they should influence the behavior of proteins and other macromolecules due to crowding, if the proteins and other macromolecules are evenly distributed within the cell. However, local concentrations of proteins may vary from one compartment to another or in proximity to membranes (24).

Besides macromolecules, abundant molecules in cells are charged metabolites and ions. According to the database BioNumbers (25), the most abundant ions in eukaryotic cells are K^+ (100 mM), Na^+ (10 mM), Ca^{2+} (10 – 100 nM), Mg^{2+} (0.5 – 10 mM) and Cl^- (5 – 100 mM). These concentrations change depending on the tissue and whether the corresponding ion is free in solution or bound to protein complexes. The concentrations of metals such as Fe, Zn, Cu or Mn are in the low mM to µM range. The most abundant metabolites in human cells are HCO_3^- (~10 mM), 2,3-diphosphoglyceric acid and glutathione (both ~ 5 mM), ATP and glutamic acid (both ~1-2 mM). Together with other metabolites such as glycerol, trimethyl-amine-N-oxide (TMAO) or different sugar compounds, the total concentration of metabolites in human cells add up to 100 – 300 mM. The presence of those co-solutes will also influence the properties of proteins in a cell. High concentrations of charged metabolites and ions change the intracellular ionic strength and affect the pH of the environment, resulting in changed conformations or stabilities of proteins (26-28). Other metabolites such as TMAO or sugars are known to have stabilizing effects on proteins (29).

The influence of steric exclusion is difficult to disentangle from nonspecific attractive interactions between crowders and the studied protein. One study found a predominant steric exclusion effect of the intracellular environment of HEK 293 cells on a hinge-protein. This crowding sensor was more compact in cells after a hypertonic shock, compared to cells under physiological conditions, which was attributed to increased crowding (30). It was also observed that the cytosol of *E. coli* was able to stabilize the small protein GB1 (31). Contrary to those findings, evidence was found that nonspecific charge interactions between the studied proteins and intracellular co-solutes dominate over steric exclusion effects in other cases (32). Reconstituted *E. coli* cytosol had a destabilizing effect on CI2 (33), and the extracellular protein VlsE was destabilized in U2OS cells (34). The influence of intracellular proteins and metabolites on the stability of proteins, mainly due to charge interactions with the surface of proteins, has been termed ‘quinary interactions’ (35, 36). Quinary interactions tune the stability of proteins by interactions of the intracellular environment with the surface of the protein. In this framework, surface exposed residues are, together with hydrophobic core residues, responsible for protein stability, depending on their environment. In summary, the inside of a cell is a complex mixture of possible interaction partners, which can either have a stabilizing or destabilizing effect on proteins, and dissecting these contributions remains an important challenge.

1.2. Intrinsically disordered proteins (IDPs)

1.2.1. Structural and dynamical aspects of IDPs

The classical view of protein folding is that the primary structure, the amino acid sequence, determines the secondary and tertiary structure of proteins, and that the function of a protein is determined by its structure (37). However, a lot of proteins have now been found to have very specific functions although they either contain unstructured regions or completely lack any secondary or tertiary structure (38-40) (Fig. 3). Proteins with this property are called intrinsically disordered proteins (IDPs).

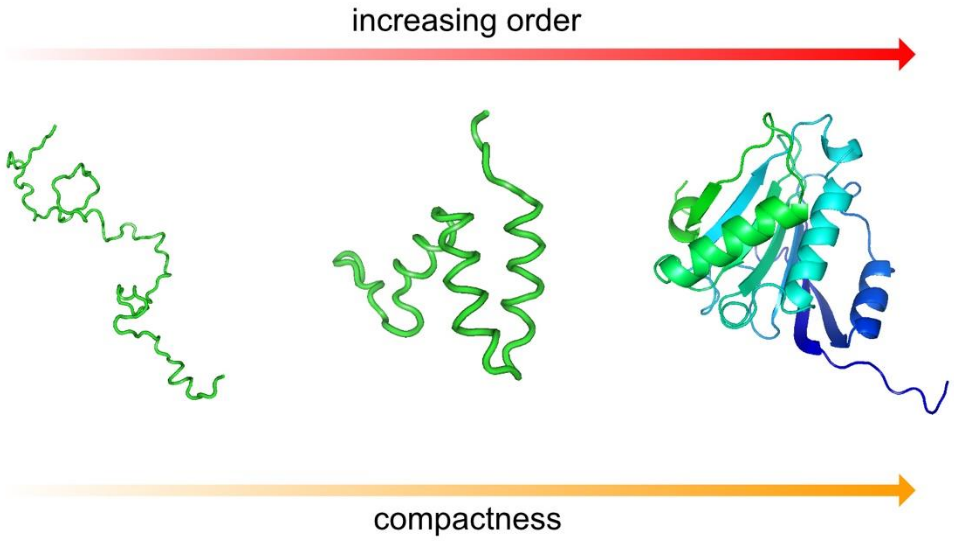


Fig. 3: Scheme of differently ordered and compacted proteins, ranging from completely unfolded and very loosely packed proteins (e.g. prothymosin α) at the left, to proteins with a higher structural content, but still very dynamic conformational ensemble (a molten globule such as NCBD) in the middle, to compact and well-ordered proteins at the right (e.g. eIF4E), where only the N-terminus is unstructured.

A characteristic property of disordered proteins is the low content of bulky hydrophobic and aromatic amino acids. These amino acids are required to form the stable hydrophobic core of folded proteins. On the other hand, polar amino acids and structure breaking amino acids are found in higher numbers (41). The content of charged amino acids and their distribution within the primary structure can have a major impact on the degree of disorder and dimensions of IDPs (27, 42). One example of an completely disordered IDP is ProTα, a small nuclear protein (43). It consists of 110 amino acids (12 kDa), of which 54 amino acid residues are negatively charged. The repulsive effect of the high spatial density of negative charges leads to a completely disordered structure (44) (Fig. 3). Only in the presence of kosmotropes such as trifluoroethanol (TFE) at acidic conditions, ProTα was found to adopt an α-helical structure.

The characteristics of IDPs can be displayed by plotting the absolute mean net charge of an IDP against its mean hydropathy (Fig. 4a) (45). Disordered proteins are mostly found to the left of the black diagonal line, whereas folded proteins are typically characterized by a lower net charge and higher mean hydropathy. Another way of representing characteristics of an IDP is by considering the fraction of negatively and positively charged residues and their distribution within the sequence (42) (Fig. 4b). The plot allows IDPs to be grouped according to their polyampholytic or –electrolytic properties. Weak polyampholytes and –electrolytes tend to have a compact and globular structure (regions 1 & 2), whereas strong polyampholytes and –electrolytes are more coil like and expanded (regions 3 - 5).

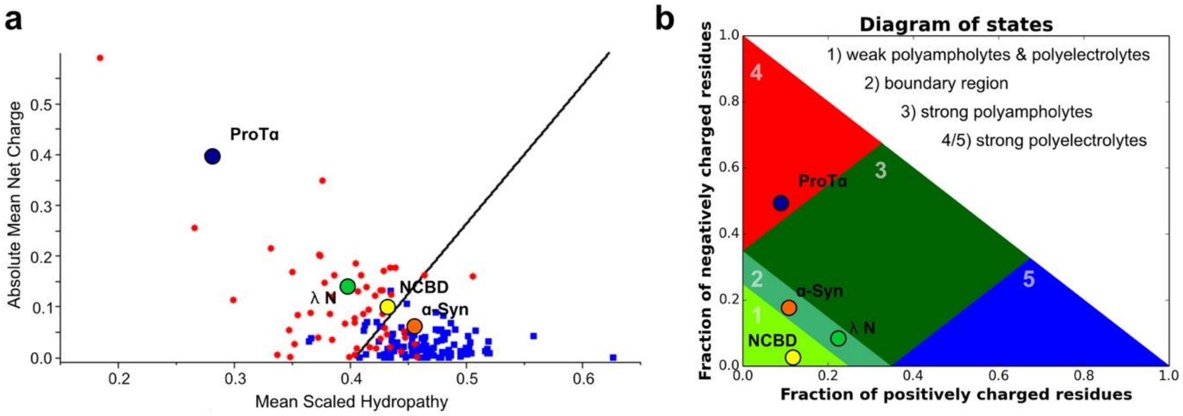


Fig. 4: Examples of how the disordered character of proteins can be displayed. (a) A plot showing the absolute mean net charge of the IDPs ProTα (dark blue), λN (green), NCBD (yellow) and α-Syn (orange) against their mean hydropathy (green) in comparison to other disordered (red) and ordered (blue) proteins. The black line describes an approximate boundary between ordered and disordered proteins (plot was generated at pondr.com) (45). (b) A diagram showing the same IDPs as in (a), but the fraction of negatively charged residues of the IDPs is plotted against the fraction of positively charged residues. The diagram is divided into different regions, according to the polyampholytic or polyelectrolytic properties of the IDPS (42).

Four different IDPs used in this thesis are shown in these plots: Prothymosin α (ProTα), an IDP with a high content of negative charges that is completely disordered (27, 44); λN, an IDP with a positive net charge and some transient and fluctuating secondary structure (46); the nuclear coactivator binding domain of CBP (NCBD), a molten globule with high α-helical content (47); and α-synuclein (α-Syn), an aggregation-prone IDP that is, in its monomeric state, completely disordered, but folds into an α-helical structure after binding to membranes (48). The Uversky plot (Fig. 4a) takes the absolute content of charged and hydrophobic residues as a measure for disorder. Three out of four IDPs (ProTα, λN, NCBD) are on the disordered side of the approximate boundary between ordered and disordered proteins, whereas α-Syn falls into the region dominated by ordered proteins. This behavior illustrates that a clear distinction between disordered and structured proteins based on charge and hydropathy alone cannot be made. The Das-Pappu plot (Fig. 4b) focuses on the effect of charge content and its distribution within the sequence on the dimensions of IDPs (42). Strong polyelectrolytes such as ProTα are characterized as a swollen coil, which represents its expanded dimensions well. α-Syn and λN are characterized as expanded chains, according to their disordered character. NCBD falls into the state of globules, in agreement with its compact molten globule-like conformations with pronounced α-helical structure (47).

Describing the structural properties of IDPs is difficult due to their dynamic nature and large degree of conformational heterogeneity. A reasonable way of representing IDPs is to show an ensemble of possible conformations, which can, e.g., be generated based on simulations combined with structural constraints from NMR or SAXS measurements. These illustrations highlight how IDPs can adopt even very expanded structures and sample their conformational space extensively (Fig. 5). Owing to their disordered nature, IDPs can also be assumed to have similar properties as unfolded proteins. This connection is of particular interest because unfolded states are difficult to investigate without destabilizing proteins with denaturants or by amino acid exchanges.

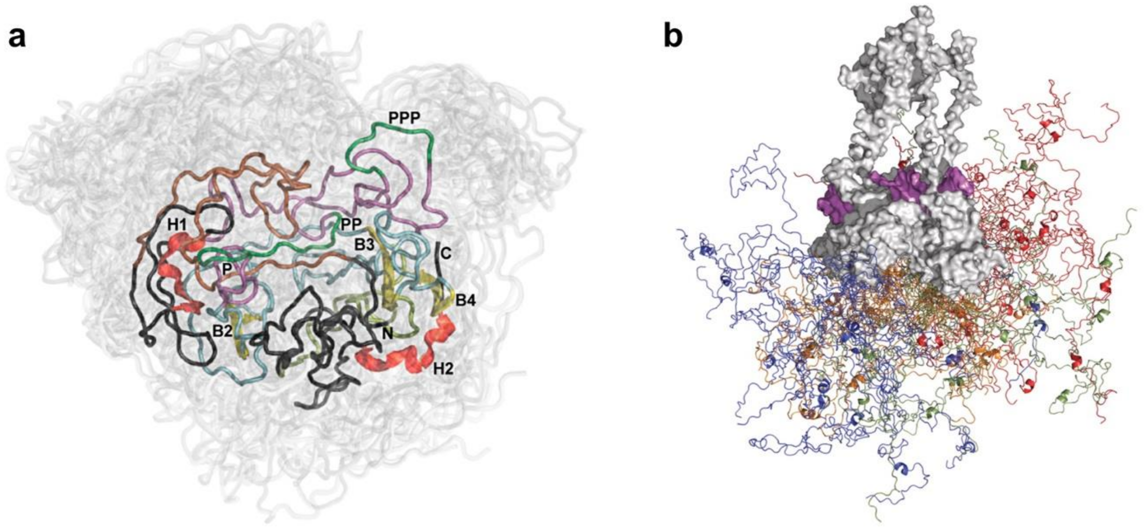


Fig. 5: Representation of structural ensembles of IDPs. (a) The structure of Alzheimer's related IDP Tau (htau40), calculated from paramagnetic relaxation enhancement (PRE) data, where the colored structure is a snapshot of the ensemble, 20 more structures are shown in grey (49). (b) A representative model of full-length p53 (structured part in grey), bound to DNA (magenta), derived from NMR and SAXS data. 20 copies of the disordered N-terminal transactivation domain TAD of each of the four monomers are shown in color (50).

Many IDPs undergo folding transitions upon binding to other proteins or after posttranslational modifications. An example of coupled binding and folding is the interaction between NCBD and activator for thyroid hormone and retinoid receptor (ACTR). ACTR is completely disordered, but folds into an α-helical structure upon binding to NCBD (51). However, NCBD, a molten globule with high secondary structure content in its free state, can also bind to interferon regulatory factor 3 (IRF-3), which forms a different structure in the bound state as in complex with ACTR (47) (Fig. 6).

Many other coupled binding and folding processes depend on posttranslational modifications of the corresponding IDPs. The transcriptional coactivator CBP binds to pKID via its KIX domain. The phosphorylation of disordered KID is necessary to increase the binding affinity to the complex. Upon binding of CBP to KIX, pKID folds to form two α-helices (52). Phosphorylation can be sufficient to induce folding of disordered proteins, as in the case of 4E-BP2, which is involved in transcription regulation. The phosphorylation of two different residues induces the formation of a four-stranded β-domain. However, not all IDPs undergo folding transitions upon interaction with other proteins. They can remain disordered after binding to their target, forming dynamic and unstructured 'fuzzy' complexes (53).

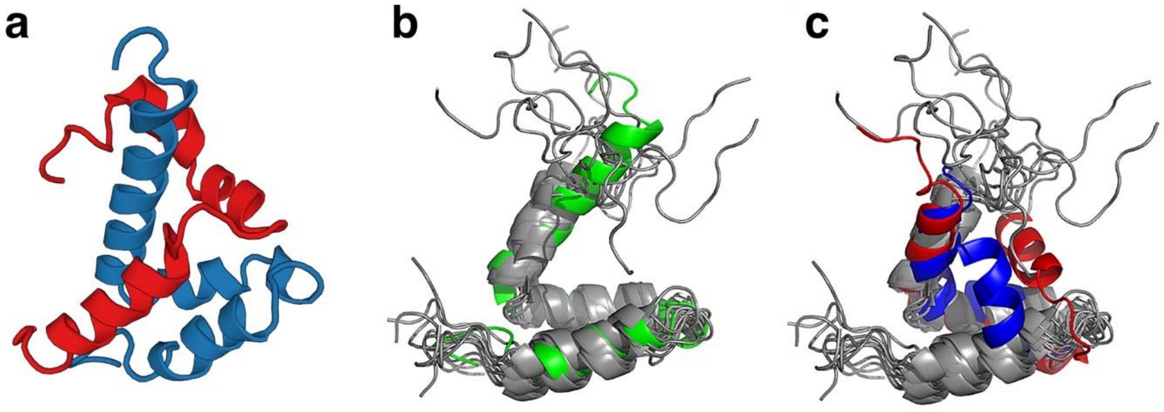


Fig. 6: NCBD in complex with ACTR or IRF-3. (a) The structure of folded ACTR (red), bound to NCBD (blue) (51). (b/c) The structural ensemble of free NCBD (grey, 10 structures), overlaid with NCBD in complex with ACTR (green) and IRF-3 (red) (47).

1.2.2. Biological roles of IDPs

IDPs may constitute up to 30% of eukaryotic proteomes (54), thus they are involved in a plethora of interactions. Proteins with disordered regions (IDRs) or with complete disorder often contain several binding motifs and are part of dynamic networks of interacting proteins in signaling and regulatory processes (55). A prominent example is the tumor suppressor p53. The N- and C-termini of p53 are unstructured (Fig. 5b), making ca. 40 % of its structure intrinsically disordered (56). p53 is the central element of a network consisting of hundreds of interactions (57), and it ensures, amongst other roles, that DNA damage is detected. In unstressed cells, p53 occurs at low concentration and is downregulated by Mdm2 (58), which binds to the unstructured N-terminal TAD domain of p53 and blocks its function. Upon cellular stress, specific kinases phosphorylate the N-terminus of p53, leading to the release of Mdm2 and enhancing the binding of the N-terminus to several domains of the transcriptional coactivators CBP and p300 (59).

1.2.3. Effect of crowding on IDPs

The free energy surfaces of IDPs are relatively flat and featureless compared to folded proteins (60). This property is believed to make them more susceptible towards interactions with their environment, such as to macromolecular crowding. Simulations showed that crowding leads to a compaction of a coarse-grained model of an IDP (61) and of a self-avoiding polymer (62). This is in accordance with the assumption that compact conformations are favored under crowded conditions. The compaction of disordered or unfolded proteins was experimentally found in several cases using crowders such as dextrans or Ficoll (63-65). The same studies also concluded that the investigated IDPs remained unstructured and highly dynamic under crowded conditions. There were also cases reported where secondary structure was induced in IDPs under crowded conditions (66). However, the induction of secondary structure was found in the presence of glucose, which is rather a co-solute than a macromolecular crowder. Thus, the gain in secondary structure can be attributed to attractive interactions.

1.3. Methods to study intrinsically disordered proteins in living cells

1.3.1. Studying proteins in the test tube and in living cells

Most of what is known about protein structure and function was found by studies performed in the test tube (*in vitro*) under chemically defined conditions in aqueous buffers. Studying proteins in their natural environment such as in living cells (*in vivo*) or embedded in an extracellular matrix is challenging because of the complexity of its surrounding. However, recent advances have made it possible to study folded proteins and especially IDPs also in living cells. Some of these methods are outlined below.

Experimental and computational methods to study protein structure and function under *in vitro* conditions were constantly developed and improved during the last decades. Spectroscopic methods, such as UV absorption, circular dichroism (CD), small-angle X-ray (SAXS) and neutron scattering (SANS) or fluorescence spectroscopy are widely used to analyze the structures of pure proteins in solution. SAXS and SANS were shown to be useful to obtain information about dimensions and shapes of ordered and disordered proteins (67). Both methods are demanding in sample preparation, because large amounts of pure protein are needed. CD allows the presence or absence of secondary structure elements in a protein to be determined by measuring the absorption of circularly polarized light. The addition of kosmotropic trimethylamine-N-oxide (TMAO) to a solution of the disordered fragment of human glucocorticoid receptor (GR 1-500), e.g., induced secondary structure, as shown by the differences in the CD spectra (68) (Fig. 7a).

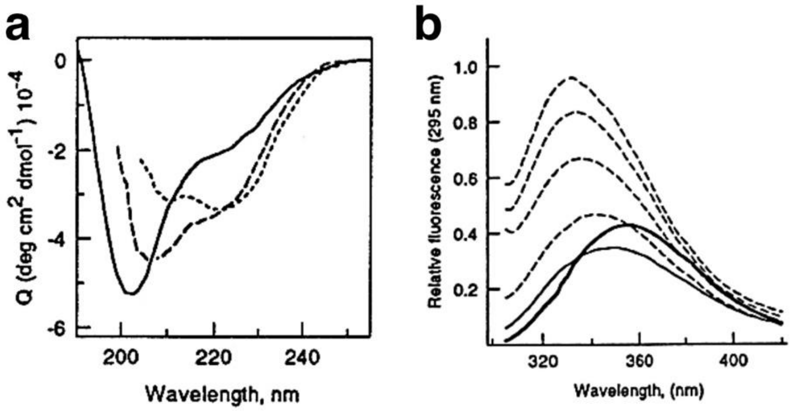


Fig. 7: Characterization of an IDP with CD and fluorescence spectroscopy. (a) CD spectra of GR 1–500 in the absence of TMAO (solid line) and presence of TMAO (dashed and dotted line). (b) Fluorescence emission spectra (295 nm excitation) of native GR 1–500 (thin solid lines), in 7 M guanidinium chloride (bold solid lines), and in the presence of TMAO (dashed lines) (68).

Complementary to CD, ensemble fluorescence spectroscopy can be used to study the structure of proteins. The fluorescence emission of tryptophan is very sensitive to its local environment and can therefore be used to analyze whether a protein is folded, unfolded or changes its compactness. The compaction and folding of disordered GR 1-500, induced by TMAO, is shown by a shift of the fluorescence emission maxima to shorter wavelengths (Fig. 7b), compared to natively disordered or denatured GR 1-500 (thin and bold solid line) (68). Unlike CD spectroscopy, fluorescence spectroscopy is well suited to study proteins in living cells. Using fluorescently tagged proteins, it is possible to specifically target the protein of interest even in the complex environment of a cell. Methods such as fluorescence recovery after photobleaching (FRAP) have been used to analyze diffusion of fluorescently labeled dextrans and Ficolls in MDCK and Swiss 3T3 cells (69) or proteins in *E. coli* (70) and eukaryotic cells (71, 72).

A versatile method to study protein dynamics is fluorescence correlation spectroscopy (FCS), where fluorescence intensity fluctuations are analyzed. FCS can be used to observe dynamic processes from seconds to nanoseconds (73), which includes, besides translational diffusion, also protein chain reconfiguration (74). FCS is thus well suited to study the influence of the local environment on protein diffusion and chain reconfiguration in living cells. Until recently, the application of FCS in living cells mainly focused on the translational diffusion of fluorescently labeled proteins (75, 76) or the diffusion of proteins or lipids incorporated into cell membranes (77). As an extension of FCS, fluorescence image correlation methods allow the diffusion of plasma-membrane proteins to be studied in living cells (78, 79).

Dimensions of proteins can be measured by Förster resonance energy transfer (FRET) spectroscopy, where distances between two fluorescent labels of up to 10 nm are measured (80, 81). Upon excitation of a donor fluorophore, which can be a fluorescent protein or a small organic dye, a distance dependent, non-radiative energy transfer to the acceptor fluorophore occurs. The transfer efficiency is a measure for the interdye distance. It is calculated from the ratio of emitted acceptor photons over the sum of emitted donor and acceptor photons, where a high transfer efficiency indicates a short interdye distance. This approach was used in living cells to study the protein folding kinetics and stabilities of PGK, labeled with two fluorescent proteins as FRET sensors (82). After destabilizing the intracellular protein with laser-induced temperature jump, the reequillibration of the FRET signal was analyzed. It was found that PGK was more stable than *in vitro*, but the folding and unfolding kinetics were only marginally affected by the cell (83).

Super-resolution microscopy has become increasingly popular because it allows fluorescent probes to be imaged below the diffraction limit (84). With spatial resolutions of less than 20 nm, super-resolution microscopy is mainly used for the localization of proteins or protein complexes such as nuclear pores in the nuclear envelope (85) or vesicles decorated with labeled α-synuclein (86). However, the temporal resolution of super-resolution methods is currently limited to 0.1 – 0.5 s (87), which does not allow diffusional or fast conformational dynamics of proteins to be studied.

To solve the structure of proteins on an atomic level, mostly X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy are used. X-ray crystallography is very powerful in resolving structures in the low Ångström range. However, a major limitation is that disordered proteins do not allow defined X-ray diffraction patterns to be recorded and that only static structural information are obtained. It was shown that NMR can cope with those problems. An increasing number of protein structures deposited in the PDB database were solved by NMR, and NMR is a key method to elucidate structural properties of disordered proteins *in vitro* and in living cells (88).
α-Synuclein is an IDP that was extensively studied by NMR both *in vitro* as well as *in vivo*. A combination of NMR spectroscopy and ensemble molecular dynamics (MD) was used to map the topology of α-synuclein (89). A very compact structural ensemble was found with long-range contacts between the C-terminal tail and the

aggregation promoting hydrophobic middle part (NAC region). These contacts have been suggested to shield the NAC region and prevent α-synuclein from forming amyloid fibrils (89). Other NMR experiments on α-synuclein studied its amyloid fibril structure (90) and the interaction of the IDP with lipids and membranes (91). The potential of NMR as a method for in-cell experiments was shown by studying α-synuclein in living neuronal cells (92), where it was found that the protein remained disordered within living cells. NMR was also used to resolve interactions between proteins and its surrounding in cell lysates (31) and living cells (93) with residue-level resolution.

1.3.2. Practical considerations for in-cell studies

The simplest way to study proteins in a cell-like environment is by dissolving the proteins in bacterial or eukaryotic cell lysates (31, 94). This method gives insights into possible interactions of the targeted protein with intracellular ligands, but it lacks the spatial organization of intact cells. For studies employing fluorescence spectroscopy, fusion constructs of the protein of interest tagged with fluorescent proteins were co-expressed in U2OS (human osteosarcoma) and HeLa (human cervical carcinoma) cells (82) or in *E. coli* and HEK293 (human embryonic kidney) cells (30). This method has the advantage that transfection methods for expressing exogenous proteins are well established and that expressing cells can easily be identified by fluorescence microscopy. However, the use of fluorescent proteins is of limited use owing to their large size, the low brightness and photostability (95), the tendency to oligomerize (96) and the complication of controlling expression levels (97).

It is evident that not every experiment can rely on fusion proteins expressed by the cell itself. NMR, e.g. needs isotope-labeled proteins, which is mostly synthesized in *E. coli* cells, and only to a minor extent in eukaryotic cells (98). The same problems apply to techniques such as fluorescence spectroscopy, which rely on small organic dyes because of their superior photophysical properties. Common ways of labeling proteins *in vitro* with organic fluorescent dyes are by coupling functionalized dyes via maleimides to a cysteine residue or via succinimidyl esters to an amino residue. The targeted residues may either occur naturally or are introduced by site-specific mutagenesis. This methodology has successfully been used on many folded and disordered proteins (99). Recently, the use of click chemistry together with the site-specific introduction of unnatural amino acids has also been shown to yield satisfying results (100). Labeling proteins in living cells is a challenging task, especially because of the required target specificity and the remaining fluorescent background. Advances for in-cell labeling were made by developing membrane-permeable dyes which exhibit their fluorogenic character only after binding to its target (101). So far, this approach has been useful in combination with super-resolution microscopy, but less for structural studies of proteins.

Several methods have been established to deliver labeled macromolecules into living cells. A rather crude approach is ‘scrape-loading’, where cells are loaded with labeled molecules by manually disrupting the cell membranes in the presence of the sample (102). Electroporation was shown to be a more efficient way of loading cells with labeled proteins. Micromolar concentrations of protein were successfully delivered into *E. coli* cells (103) and into several mammalian cells, such as HeLa or SK-N-SH (human neuroblastoma) cells (92), without affecting the viability of the cells drastically. Other loading methods, such as import of proteins using cell penetrating peptides (CPPs) (104) or pore forming toxins (105) were shown to work in principle. However, especially CPPs can be difficult to work with because the cargo can be trapped in endosomes with only little escape of the imported proteins into the cytosol (106).

Microinjection is a robust and reproducible tool to transfect living cells with DNA. This technique has also been used to inject other macromolecules into the cytosol or nuclei of cells, while maintaining a high survival rate of cells. For instance, the intracellular diffusion of fluorescently labeled dextrans and Ficolls in MDCK (canine kidney) and Swiss 3T3 (murine fibroblast) cells was studied after delivering the sample by microinjection (69). Fluorescently labeled proteins such as IgG or β-galactosidase were microinjected into cultured murine muscle cells (71), showing the versatile usage of microinjection. Although microinjection is not a high-throughput method, it still allowed enough *X. laevis* oocytes to be injected with the protein GB1 to perform in-cell NMR experiments (107).

1.3.3. Single-molecule FRET spectroscopy

Single-molecule spectroscopy in combination with Förster resonance energy transfer (FRET) is a powerful tool to study structures and dynamics of proteins (73). The key feature of single-molecule FRET spectroscopy is that proteins are studied individually, without averaging over an ensemble of molecules. Several aspects of proteins can be studied, such as distance distributions within a protein, folding kinetics, or dynamic processes such as diffusion and fast protein reconfigurations.

The readout of a single-molecule FRET measurement can be displayed as a trace of photons (Fig. 8a), where every passage of a single FRET-labeled protein through the confocal volume of the microscope generates a burst of green and red photons. The subsequent data analysis allows to determine protein dimensions and dynamics.

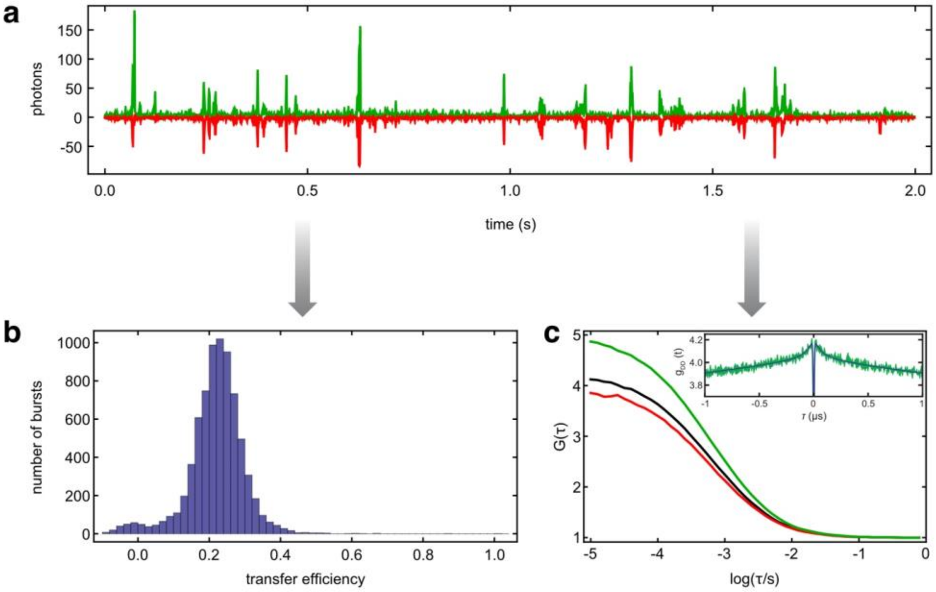


Fig. 8: Data analysis of single-molecule FRET data. (a) Photon trace showing fluorescence bursts in the donor (green) and acceptor (red) channel. Each burst originates from one FRET-labeled molecule passing the observation volume. (b) Transfer efficiency histogram calculated from the detected photon bursts. (c) Fluorescence intensity correlations (donor-donor in green, acceptor-acceptor in red, donor-acceptor in black) showing amplitudes in the millisecond range from translational diffusion or in the nanosecond range (inset) originating from chain reconfiguration dynamics.

Protein dimensions

As described elsewhere (99, 108), the ratio of detected acceptor (n_A) and donor photons (n_D) is used to calculate the transfer efficiency ($E = n_A/(n_A + n_D)$), which is a measure for the distance between the two dyes. The transfer efficiency is calculated for all the detected bursts, upon which a transfer efficiency histogram is built (Fig. 8b). Since this is a single-molecule technique, the number of populations in the transfer efficiency histogram report on the number of states the observed protein can adopt, assuming the populations interconvert slower than their diffusion time through the observation volume (109). A fit of those populations allows the mean transfer efficiency, $\langle E \rangle$, to be determined, from which quantities such as mean interdye distances or radii of gyration can be inferred.

The mean transfer efficiency $\langle E \rangle$ can be expressed as a function of the distance dependence of the transfer efficiency $E(r)$, weighted by a dye-to-dye distance distribution $P(r)$ which is sampled by the chain,

$$\langle E \rangle = \int_a^{l_c} E(r)P(r)dr$$

with

$$E(r) = \frac{1}{1 + (r/R_0)^6}$$

where a is the distance of the closest approach of the dyes, l_c is the contour length of the protein segment between the two dyes, and R_0 is the Förster radius of the chosen dye pair. For the distribution of distances $P(r)$, simple polymer models, such as a Gaussian chain are often used (110, 111),

$$P(r) = 4\pi r^2 \left(\frac{3}{2\pi\langle r^2 \rangle}\right)^{3/2} exp\left(-\frac{3r^2}{2\langle r^2 \rangle}\right)$$

where $\langle r^2 \rangle$ is the mean squared interdye distance. Experimentally accessible mean interdye distances range from approximately 2 – 10 nm, given the Förster radius of the dye pair used for single-molecule FRET experiments. Assuming a Gaussian chain, $\langle r^2 \rangle$ can be converted into the mean radius of gyration, R_g, using $R_g^2 = \langle r^2 \rangle/6$.

Protein dynamics

FCS

By correlating the fluorescence intensity fluctuations of a single-molecule FRET measurement (Fig. 8c), information about dynamic processes, such as translational diffusion or triplet state dynamics, can be obtained (112). The correlations are carried out according to

$$g_{ij}(\tau) = \frac{\langle n_i(t)n_j(t+\tau)\rangle_t}{\langle n_i(t)\rangle_t \langle n_j(t)\rangle_t}$$

with $i, j = A, D$ (A = acceptor, D = donor). $n_i(t)$ and $n_j(t+\tau)$ are the fluorescence count rates in the detector channels i and j and $\langle \ldots \rangle_t$ denotes the time average of the signal. The obtained cross-correlation curves of the donor and the acceptor channels are fitted with

$$g_{AD}(\tau) = 1 + \frac{1}{N_{AD}} \frac{1 + c_T^{(AD)} e^{-\tau/\tau_T^{(AD)}}}{(1 + \tau/\tau_D)(1 + \tau/s^2\tau_D)^{1/2}}$$

where c_T and τ_T describe the amplitude and the decay time of the triplet state, respectively. N_{AD} describes the average number of donor-acceptor labeled proteins in the confocal volume, s the ratio of axial over lateral radius of the confocal volume and τ_D the translational diffusion time of the labeled protein through the confocal volume of the microscope.

Nanosecond FCS (nsFCS)

Chain dynamics of IDPs occur on a timescale from approximately 20 to 200 ns with a segment length of ~30 to ~200 amino acid residues and can be assessed with nsFCS (113). As earlier described (74, 114), the donor and acceptor autocorrelation curves and the donor-acceptor cross-correlation curve are calculated and fitted globally up to a delay time of 1 µs using

$$g_{i,j}(\tau) = 1 + \frac{1}{N}\left(1 - c_{AB} e^{-|t-t_0|/\tau_{AB}}\right)\left(1 + c_{CD} e^{-|t-t_0|/\tau_{CD}}\right)\left(1 + c_T e^{-|t-t_0|/\tau_T}\right)$$

with $i, j = A, D$ (A = acceptor, D = donor). N corresponds to the average number of proteins in the confocal volume and the three multiplicative terms describe photon antibunching (AB), chain dynamics (CD), and triplet blinking (T) of the dyes. The decay time of the chain dynamics (τ_{CD}) is a shared fit parameter and is extracted from the global fit of the three correlations. After correction of τ_{CD} for the transfer efficiency of the labeled protein (115), the chain reconfiguration dynamics of a protein can be analyzed.

1.3.4. Single-molecule FRET spectroscopy in living cells

Single-molecule FRET spectroscopy has been used extensively to study proteins *in vitro*, e.g. how dimensions of IDPs depend on the ionic strength of the solution (27). Applying single-molecule FRET spectroscopy to living cells is a challenging task which only few research groups tried to approach yet, but it promises to increase the knowledge about how proteins function in living cells.

First in-cell single-molecule FRET experiments were performed by delivering labeled DNA and proteins into *E. coli* and *S. cerevisiae* cells using electroporation (103). The samples were successfully internalized and single molecules were tracked in live bacteria. Even FRET of intact molecules in living bacteria was measured to analyze the conformations of the observed species. At the time of the start of this thesis, only one publication existed where the successful microinjection and detection of FRET-labeled proteins in live mammalian cells on a single-molecule level was reported (116).

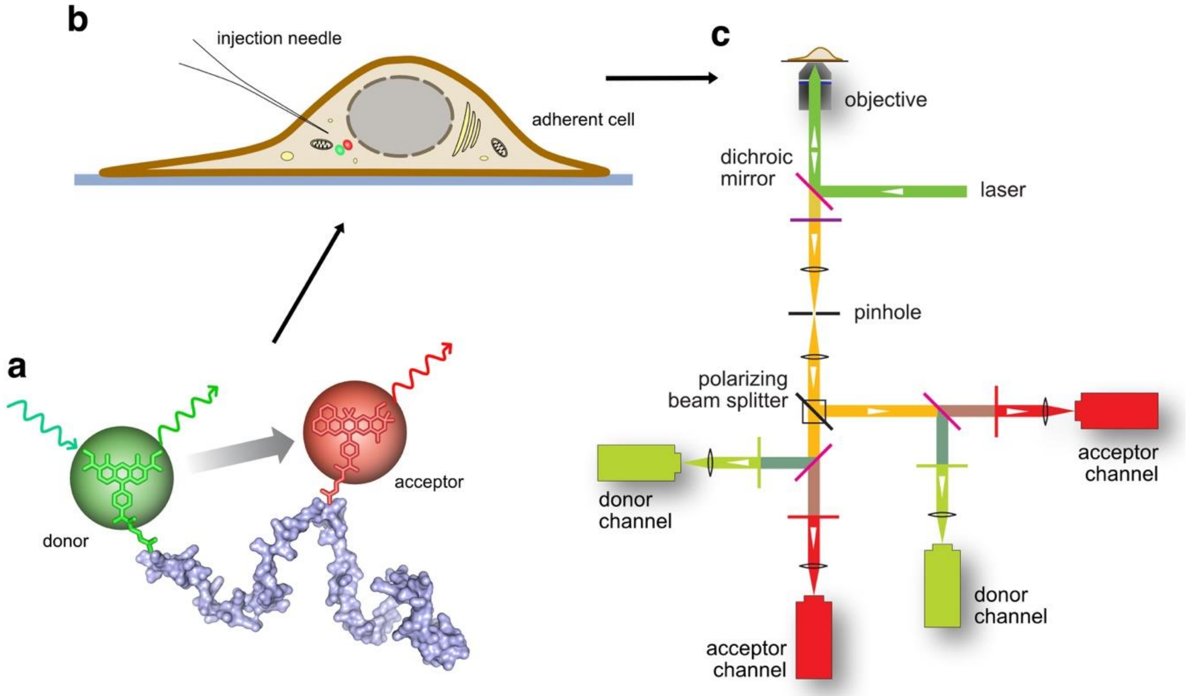


Fig. 9: Overview of the process for analyzing FRET-labeled proteins in living cells. (a) FRET-labeled ProTα, indicating FRET from the donor (green) to the acceptor dye (red). (b) Microinjection of labeled protein into adherent HeLa cells. (c) Scheme of the confocal microscope used for single-molecule FRET measurements in cells.

Although single-molecule FRET is measured with a confocal microscope to minimize background fluorescence, intracellular autofluorescence is still a major problem for in-cell single-molecule FRET spectroscopy. Cellular autofluorescence is mainly caused by molecules such as NADH or flavins (117), with fluorescence emission up to a wavelength of ~520 nm. Many fluorescent dyes used for FRET (Fig. 9a) absorb and emit light below or in the range where autofluorescence occurs (e.g. Alexa 488 and Alexa 594). The problem of autofluorescence can thus be reduced by choosing dyes which are shifted towards larger emission wavelengths.

Choosing the best dye pair depends on the problem to be solved. However, special care has to be taken that the dyes don’t interfere with the experiment. Especially hydrophobic dyes can non-specifically bind to, e.g., cellular membranes (118, 119) or mitochondria (120). Based on previous publications (107, 116) and on many tests with other methods such as scrape-loading (102), cell penetrating peptides (104) and electroporation (93), microinjection was chosen to deliver the FRET-labeled proteins into living cells (Fig. 9b). Microinjection should allow pM to nM concentrations of labeled proteins to be delivered, which is needed for single-molecule spectroscopy. The single-molecule FRET measurements are carried out using a confocal microscope (Fig. 9c), where the labeled proteins get excited with a laser. The emitted photons from the donor and acceptor dyes are collected by a high aperture objective and split according to the color into four detection channels, where they are detected separately by single-photon avalanche detectors (99).

References

1. Minton AP, Wilf J (1981) Effect of macromolecular crowding upon the structure and function of an enzyme: glyceraldehyde-3-phosphate dehydrogenase. *Biochemistry* 20(17):4821–4826.
2. Zhou H-X, Rivas G, Minton AP (2008) Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu Rev Biophys* 37(1):375–397.
3. Minton AP (2000) Effect of a concentrated “inert” macromolecular cosolute on the stability of a globular protein with respect to denaturation by heat and by chaotropes: A statistical-thermodynamical model. *Biophys J* 78(1):101–109.
4. Hayashi K, Nakazawa M, Ishizaki Y, Hiraoka N, Obayashi A (1985) Stimulation of intermolecular ligation with E. coli DNA ligase by high concentrations of monovalent cations in polyethylene glycol solutions. *Nucleic Acids Res* 13(22):7979–7992.
5. Behe MJ, Englander SW (1978) Sickle hemoglobin gelation. Reaction order and critical nucleus size. *Biophys J* 23(1):129–145.
6. Mittal S, Singh LR (2013) Denatured state structural property determines protein stabilization by macromolecular crowding: a thermodynamic and structural approach. *PLOS ONE* 8(11):e78936.
7. Dhar A, et al. (2010) Structure, function, and folding of phosphoglycerate kinase are strongly perturbed by macromolecular crowding. *Proc Natl Acad Sci USA* 107(41):17586–17591.
8. Ådén J, Wittung-Stafshede P (2014) Folding of an unfolded protein by macromolecular crowding in vitro. *Biochemistry* 53(14):2271–2277.
9. Batra J, Xu K, Qin S, Zhou H-X (2009) Effect of Macromolecular Crowding on Protein Binding Stability: Modest Stabilization and Significant Biological Consequences. *Biophys J* 97(3):906–911.
10. Hatters DM, Minton AP, Howlett GJ (2002) Macromolecular crowding accelerates amyloid formation by human apolipoprotein C-II. *J Biol Chem* 277(10):7824–7830.
11. Yamin G, et al. (2005) Forcing nonamyloidogenic beta-synuclein to fibrillate. *Biochemistry* 44(25):9096–9107.
12. Ittah V, Kahana E, Amir D, Haas E (2004) Applications of time-resolved resonance energy transfer measurements in studies of the molecular crowding effect. *J Mol Recognit* 17(5):448–455.
13. Tokuriki N, et al. (2004) Protein folding by the effects of macromolecular crowding. *Protein Sci* 13(1):125–133.
14. Kuttner YY, Kozer N, Segal E, Schreiber G, Haran G (2005) Separating the contribution of translational and rotational diffusion to protein association. *J Am Chem Soc* 127(43):15138–15144.
15. Dix JA, Verkman AS (2008) Crowding effects on diffusion in solutions and cells. *Annu Rev Biophys* 37(1):247–263.
16. Wang Y, Benton LA, Singh V, Pielak GJ (2012) Disordered Protein Diffusion

under Crowded Conditions. *J Phys Chem Lett* 3(18):2703–2706.

17. Crowley PB, Brett K, Muldoon J (2008) NMR spectroscopy reveals cytochrome c-poly(ethylene glycol) interactions. *ChemBioChem* 9(5):685–688.
18. Miklos AC, Sarkar M, Wang Y, Pielak GJ (2011) Protein crowding tunes protein stability. *J Am Chem Soc* 133(18):7116–7120.
19. McGuffee SR, Elcock AH (2010) Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLOS Comput Biol* 6(3):e1000694.
20. Gershenson A, Gierasch LM (2011) Protein folding in the cell: challenges and progress. *Curr Opin Struc Biol* 21(1):32–41.
21. Zimmerman SB, Trach SO (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. *J Mol Biol* 222(3):599–620.
22. Conlon I, Raff M (2003) Differences in the way a mammalian cell and yeast cells coordinate cell growth and cell-cycle progression. *J Biol* 2(1). doi:10.1186/1475-4924-2-7.
23. Cheung MC, et al. (2013) Intracellular protein and nucleic acid measured in eight cell types using deep-ultraviolet mass mapping. *Cytom Part A* 83(6):540–551.
24. Spitzer J, Poolman B (2013) How crowded is the prokaryotic cytoplasm? *FEBS Lett* 587(14):2094–2098.
25. Milo R, Jorgensen P, Moran U, Weber G, Springer M (2010) BioNumbers—the database of key numbers in molecular and cell biology. *Nucleic Acids Res* 38(suppl_1):D750–D753.
26. Uversky VN, et al. (2000) Zn(2+)-mediated structure formation and compaction of the “natively unfolded” human prothymosin alpha. *Biochem Biophys Res Commun* 267(2):663–668.
27. Müller-Späth S, et al. (2010) Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc Natl Acad Sci USA* 107(33):14609–14614.
28. Di Russo NV, Estrin DA, Martí MA, Roitberg AE (2012) pH-Dependent conformational changes in proteins and their effect on experimental pK(a)s: the case of Nitrophorin 4. *PLOS Comput Biol* 8(11):e1002761.
29. Cho SS, Reddy G, Straub JE, Thirumalai D (2011) Entropic stabilization of proteins by TMAO. *J Phys Chem B* 115(45):13401–13407.
30. Boersma AJ, Zuhorn IS, Poolman B (2015) A sensor for quantification of macromolecular crowding in living cells. *Nat Methods* 12(3):227–229.
31. Monteith WB, Pielak GJ (2014) Residue level quantification of protein stability in living cells. *Proc Natl Acad Sci USA* 111(31):11335–11340.
32. Sarkar M, Lu J, Pielak GJ (2014) Protein crowder charge and protein stability. *Biochemistry* 53(10):1601–1606.
33. Sarkar M, Smith AE, Pielak GJ (2013) Impact of reconstituted cytosol on protein stability. *Proc Natl Acad Sci USA* 110(48):19342–19347.

34. Guzman I, Gelman H, Tai J, Gruebele M (2014) The extracellular protein VlsE is destabilized inside cells. *J Mol Biol* 426(1):11–20.

35. McConkey EH (1982) Molecular evolution, intracellular organization, and the quinary structure of proteins. *Proc Natl Acad Sci USA* 79(10):3236–3240.

36. Monteith WB, Cohen RD, Smith AE, Guzman-Cisneros E, Pielak GJ (2015) Quinary structure modulates protein stability in cells. *Proc Natl Acad Sci USA* 112(6):1739–1742.

37. Anfinsen CB (1973) Principles That Govern Folding of Protein Chains. *Science* 181(4096):223–230.

38. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6(3):197–208.

39. Uversky VN, Oldfield CJ, Dunker AK (2008) Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu Rev Biophys* 37(1):215–246.

40. Piovesan D, et al. (2017) DisProt 7.0: a major update of the database of disordered proteins. *Nucleic Acids Res* 45(D1):D1123–D1124.

41. Romero P, et al. (2001) Sequence complexity of disordered protein. *Proteins* 42(1):38–48.

42. Das RK, Pappu RV (2013) Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc Natl Acad Sci USA* 110(33):13392–13397.

43. Gomez-Marquez J, Segade F (1988) Prothymosin α is a nuclear protein. *FEBS Lett* 226(2):217–219.

44. Gast K, Damaschun H, Eckert K, Schulze-Forster K (1995) Prothymosin. alpha.: A Biologically Active Protein with Random Coil Conformation. *Biochemistry* 34(40):13211–13218.

45. Uversky VN, Gillespie JR, Fink AL (2000) Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins* 41(3):415–427.

46. Van Gilst MR, Rees WA, Das A, Hippel von PH (1997) Complexes of N Antitermination Protein of Phage λ with Specific and Nonspecific RNA Target Sites on the Nascent Transcript†. *Biochemistry* 36(6):1514–1524.

47. Kjaergaard M, Teilum K, Poulsen FM (2010) Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *Proc Natl Acad Sci USA* 107(28):12535–12540.

48. Eliezer D, Kutluay E, Bussell R, Browne G (2001) Conformational properties of alpha-synuclein in its free and lipid-associated states. *J Mol Biol* 307(4):1061–1073.

49. Mukrasch MD, et al. (2009) Structural polymorphism of 441-residue tau at single residue resolution. *PLOS Biol* 7(2):e34.

50. Wells M, et al. (2008) Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *Proc Natl Acad Sci USA* 105(15):5762–5767.

51. Demarest SJ, et al. (2002) Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* 415(6871):549–553.

52. Sugase K, Dyson HJ, Wright PE (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447(7147):1021–1025.

53. Tompa P, Fuxreiter M (2008) Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends in Biochemical Sciences* 33(1):2–8.

54. Pancsa R, Tompa P (2012) Structural Disorder in Eukaryotes. *PLOS ONE* 7(4):e34687.

55. Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN (2005) Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J* 272(20):5129–5148.

56. Römer L, Klein C, Dehner A, Kessler H, Buchner J (2006) p53--a natural cancer killer: structural insights and therapeutic concepts. *Angew Chem Int Ed Engl* 45(39):6440–6460.

57. Chatr-Aryamontri A, et al. (2017) The BioGRID interaction database: 2017 update. *Nucleic Acids Res* 45(D1):D369–D379.

58. Moll UM, Petrenko O (2003) The MDM2-p53 Interaction. *Mol Cancer Res* 1(14):1001–1008.

59. Lee CW, Martinez-Yamout MA, Dyson HJ, Wright PE (2010) Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of CREB binding protein. *Biochemistry* 49(46):9964–9971.

60. Brucale M, Schuler B, Samorì B (2014) Single-molecule studies of intrinsically disordered proteins. *Chem Rev* 114(6):3281–3317.

61. Qin S, Zhou H-X (2013) Effects of Macromolecular Crowding on the Conformational Ensembles of Disordered Proteins. *J Phys Chem Lett* 4(20):3429–3434.

62. Kang H, Pincus PA, Hyeon C, Thirumalai D (2015) Effects of Macromolecular Crowding on the Collapse of Biopolymers. *Phys Rev Lett* 114(6). doi:10.1103/PhysRevLett.114.068303.

63. Flaugh SL, Lumb KJ (2001) Effects of macromolecular crowding on the intrinsically disordered proteins c-Fos and p27(Kip1). *Biomacromolecules* 2(2):538–540.

64. Szasz C, et al. (2011) Protein Disorder Prevails under Crowded Conditions. *Biochemistry* 50(26):5834–5844.

65. Cino EA, Karttunen M, Choy W-Y (2012) Effects of molecular crowding on the dynamics of intrinsically disordered proteins. *PLOS ONE* 7(11):e49876.

66. Dedmon MM, Patel CN, Young GB, Pielak GJ (2002) FlgM gains structure in living cells. *Proc Natl Acad Sci USA* 99(20):12681–12684.

67. Goldenberg DP, Argyle B (2014) Self crowding of globular proteins studied by small-angle x-ray scattering. *Biophys J* 106(4):895–904.

68. Baskakov IV, et al. (1999) Trimethylamine N-Oxide-induced Cooperative Folding of an Intrinsically Unfolded Transcription-activating Fragment of Human Glucocorticoid Receptor. *J Biol Chem* 274(16):10693–10696.

69. Seksek O, Biwersi J, Verkman AS (1997) Translational diffusion of macromolecule-sized solutes in cytoplasm and nucleus. *J Cell Biol*

138(1):131–142.

70. Kumar M, Mommer MS, Sourjik V (2010) Mobility of cytoplasmic, membrane, and DNA-binding proteins in Escherichia coli. *Biophys J* 98(4):552–559.

71. Arrio-Dupont M, Foucault G, Vacher M, Devaux PF, Cribier S (2000) Translational Diffusion of Globular Proteins in the Cytoplasm of Cultured Muscle Cells. *Biophys J* 78(2):901–907.

72. Kuhn T, et al. (2011) Protein Diffusion in Mammalian Cell Cytoplasm. *PLOS ONE* 6(8):e22962.

73. Schuler B (2013) Single-molecule FRET of protein structure and dynamics - a primer. *J Nanobiotechnology* 11(Suppl 1):S2.

74. Nettels D, Gopich IV, Hoffmann A, Schuler B (2007) Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc Natl Acad Sci USA* 104(8):2655–2660.

75. Kim SA, Heinze KG, Schwille P (2007) Fluorescence correlation spectroscopy in living cells. *Nat Methods* 4(11):963–973.

76. Mütze J, Ohrt T, Schwille P (2011) Fluorescence correlation spectroscopy in vivo. *Laser & photonics reviews* 5(1):52–67.

77. Schwille P, Korlach J, Webb WW (1999) Fluorescence correlation spectroscopy with single-molecule sensitivity on cell and model membranes. *Cytometry* 36(3):176–182.

78. Digman MA, Gratton E (2012) Scanning image correlation spectroscopy. *Bioessays* 34(5):377–385.

79. Di Rienzo C, Gratton E, Beltram F, Cardarelli F (2013) Fast spatiotemporal correlation spectroscopy to determine protein lateral diffusion laws in live cell membranes. *Proc Natl Acad Sci USA* 110(30):12307–12312.

80. Förster T (1948) Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann Phys* 2(1-2):55–75.

81. Stryer L, Haugland RP (1967) Energy Transfer - a Spectroscopic Ruler. *Proc Natl Acad Sci USA* 58(2):719–726.

82. Ebbinghaus S, Dhar A, McDonald JD, Gruebele M (2010) Protein folding stability and dynamics imaged in a living cell. *Nat Methods* 7(4):319–323.

83. Guo M, Xu Y, Gruebele M (2012) Temperature dependence of protein folding kinetics in living cells. *Proc Natl Acad Sci USA* 109(44):17863–17867.

84. Hell SW (2007) Far-field optical nanoscopy. *Science* 316(5828):1153–1158.

85. Szymborska A, et al. (2013) Nuclear Pore Scaffold Structure Analyzed by Super-Resolution Microscopy and Particle Averaging. *Science* 341(6146):655–658.

86. Fakhree MAA, et al. (2016) The number of alpha-synuclein proteins per vesicle gives insights into its physiological function. *Sci Rep* 6:30658.

87. Jones SA, Shim S-H, He J, Zhuang X (2011) Fast, three-dimensional super-resolution imaging of live cells. *Nat Methods* 8(6):499–508.

88. Theillet F-X, et al. (2014) Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs). *Chem Rev* 114(13):6661–6714.

89. Dedmon MM, Lindorff-Larsen K, Christodoulou J, Vendruscolo M, Dobson CM (2005) Mapping long-range interactions in alpha-synuclein using spin-label NMR and ensemble molecular dynamics simulations. *J Am Chem Soc* 127(2):476–477.

90. Chen M, Margittai M, Chen J, Langen R (2007) Investigation of alpha-synuclein fibril structure by site-directed spin labeling. *J Biol Chem* 282(34):24970–24979.

91. Jao CC, Der-Sarkissian A, Chen J, Langen R (2004) Structure of membrane-bound alpha-synuclein studied by site-directed spin labeling. *Proc Natl Acad Sci USA* 101(22):8331–8336.

92. Theillet F-X, et al. (2016) Structural disorder of monomeric alpha-synuclein persists in mammalian cells. *Nature* 530:45–50.

93. Danielsson J, et al. (2015) Thermodynamics of protein destabilization in live cells. *Proc Natl Acad Sci USA* 112(40):12402–12407.

94. Wang Y, Li C, Pielak GJ (2010) Effects of proteins on protein diffusion. *J Am Chem Soc* 132(27):9392–9397.

95. van de Linde S, Heilemann M, Sauer M (2012) Live-Cell Super-Resolution Imaging with Synthetic Fluorophores. *Annu Rev Phys Chem* 63(1):519–540.

96. Costantini LM, Fossati M, Francolini M, Snapp EL (2012) Assessing the Tendency of Fluorescent Proteins to Oligomerize Under Physiologic Conditions. *Traffic* 13(5):643–649.

97. Raj A, van Oudenaarden A (2008) Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell* 135(2):216–226.

98. Dutta A, Saxena K, Schwalbe H, Klein-Seetharaman J (2012) Isotope labeling in mammalian cells. *Method Mol Biol* 831:55–69.

99. Schuler B, Müller-Späth S, Soranno A, Nettels D (2012) Application of confocal single-molecule FRET to intrinsically disordered proteins. *Method Mol Biol* 896(Chapter 2):21–45.

100. Milles S, et al. (2012) Click strategies for single-molecule protein fluorescence. *J Am Chem Soc* 134(11):5187–5195.

101. Lukinavičius G, et al. (2013) A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins. *Nat Chem* 5(2):132–139.

102. McNeil PL, Murphy RF, Lanni F (1984) A method for incorporating macromolecules into adherent cells. *J Cell Biol* 98(4):1556–1564.

103. Crawford R, et al. (2013) Long-lived intracellular single-molecule fluorescence using electroporated molecules. *Biophys J* 105(11):2439–2450.

104. Bekei B, et al. (2012) In-cell NMR in mammalian cells: part 1. *Method Mol Biol* 895:43–54.

105. Bekei B, Rose HM, Herzig M, Selenko P (2012) In-cell NMR in mammalian cells: part 2. *Method Mol Biol* 895:55–66.

106. Copolovici DM, Langel K, Eriste E, Langel Ü (2014) Cell-penetrating peptides: design, synthesis, and applications. *ACS Nano* 8(3):1972–1994.

107. Selenko P, Serber Z, Gadea B, Ruderman J, Wagner G (2006) Quantitative

NMR analysis of the protein G B1 domain in Xenopus laevis egg extracts and intact oocytes. *Proc Natl Acad Sci USA* 103(32):11904–11909.

108. Schuler B, Lipman EA, Eaton WA (2002) Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature* 419(6908):743–747.

109. Gopich IV, Szabo A (2007) Single-molecule FRET with diffusion and conformational dynamics. *J Phys Chem B* 111(44):12925–12932.

110. Merchant KA, Best RB, Louis JM, Gopich IV, Eaton WA (2007) Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations. *Proc Natl Acad Sci USA* 104(5):1528–1533.

111. Soranno A, et al. (2012) Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy. *Proc Natl Acad Sci USA* 109(44):17800–17806.

112. Rigler R, Elson ES (2001) *Fluorescence Correlation Spectroscopy* (Springer Berlin Heidelberg) doi:10.1007/978-3-642-59542-4.

113. Schuler B, Hofmann H (2013) Single-molecule spectroscopy of protein folding dynamics--expanding scope and timescales. *Curr Opin Struc Biol* 23(1):36–47.

114. Nettels D, Hoffmann A, Schuler B (2008) Unfolded protein and peptide dynamics investigated with single-molecule FRET and correlation spectroscopy from picoseconds to seconds. *J Phys Chem B* 112(19):6137–6146.

115. Gopich IV, Nettels D, Schuler B, Szabo A (2009) Protein dynamics from single-molecule fluorescence intensity correlation functions. *J Chem Phys* 131(9):095102.

116. Sakon JJ, Weninger KR (2010) Detecting the conformation of individual proteins in live cells. *Nat Methods* 7(3):203–205.

117. Aubin JE (1979) Autofluorescence of Viable Cultured Mammalian-Cells. *J Histochem Cytochem* 27(1):36–43.

118. Zanetti-Domingues LC, Tynan CJ, Rolfe DJ, Clarke DT, Martin-Fernandez ML (2013) Hydrophobic fluorescent probes introduce artifacts into single molecule tracking experiments due to non-specific binding. *PLOS ONE* 8(9):e74200.

119. Hughes LD, Rawle RJ, Boxer SG (2014) Choose your label wisely: water-soluble fluorophores often interact with lipid bilayers. *PLOS ONE* 9(2):e87649.

120. Pellett PA, et al. (2011) Two-color STED microscopy in living cells. *Biomed Opt Express* 2(8):2364–2371.

2 Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments

Andrea Soranno[1], Iwo König[1], Madeleine B. Borgia, Hagen Hofmann, Franziska Zosel, Daniel Nettels, and Benjamin Schuler

[1] Both authors contributed equally to this work

(Proceedings of the National Academy of Sciences of the USA, 2014)

The author prepared and carried out the single-molecule FRET measurements and analyzed the data to obtain dimensions of the analyzed proteins.

Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments

Andrea Soranno[1,2], Iwo Koenig[1], Madeleine B. Borgia, Hagen Hofmann, Franziska Zosel, Daniel Nettels, and Benjamin Schuler[2]

Department of Biochemistry, University of Zurich, 8057 Zurich, Switzerland

Edited by Peter E. Wright, The Scripps Research Institute, La Jolla, CA, and approved February 25, 2014 (received for review December 4, 2013)

Intrinsically disordered proteins (IDPs) are involved in a wide range of regulatory processes in the cell. Owing to their flexibility, their conformations are expected to be particularly sensitive to the crowded cellular environment. Here we use single-molecule Förster resonance energy transfer to quantify the effect of crowding as mimicked by commonly used biocompatible polymers. We observe a compaction of IDPs not only with increasing concentration, but also with increasing size of the crowding agents, at variance with the predictions from scaled-particle theory, the prevalent paradigm in the field. However, the observed behavior can be explained quantitatively if the polymeric nature of both the IDPs and the crowding molecules is taken into account explicitly. Our results suggest that excluded volume interactions between overlapping biopolymers and the resulting criticality of the system can be essential contributions to the physics governing the crowded cellular milieu.

single-molecule FRET | unfolded state collapse | excluded volume screening | Flory–Huggins theory

A surprisingly large number of eukaryotic proteins either contain substantial unstructured regions or are entirely unfolded under physiological conditions (1, 2). These "intrinsically disordered proteins" (IDPs) are involved in many crucial cellular processes, such as transcription, translation, and signal transduction; their functional and conformational properties are thus of great interest for a wide range of biological questions. Important advances in understanding the structures of IDPs have been made over the past decade, especially with spectroscopic techniques, e.g., NMR (3, 4), single-molecule fluorescence (5–7), and with atomistic and coarse-grained molecular simulations (8–10). In contrast with the stable folded structures we are familiar with from 50 y of structural biology, IDPs comprise highly heterogeneous and dynamic ensembles of conformations, which either lack stable tertiary structure altogether or fold only on binding their cellular targets (4). Important components of the cellular environment that affect IDPs include not only specific cellular ligands, but also pH and the concentration of salts (11, 12). An additional contribution that has been difficult to investigate experimentally comes from the large number of different solutes present in a cell that do not interact with an IDP specifically, but result in an environment that is densely filled with macromolecules and metabolites (12–14). Given their lack of persistent structure, the conformations of IDPs are expected to be particularly sensitive to the effects of such molecular crowding. Indeed, first experiments indicate that some IDPs gain structure upon crowding (15), whereas others do not (16–18), but may change their dimensions (19–21). The question of how the conformational distributions of IDPs respond to crowded environments is of particular current interest because IDPs have a vital role in cellular compartments and regions with very high local concentrations of proteins and RNA, such as RNA granules and nuclear pore complexes (22–25). However, a quantitative comprehension of how the concentrations and sizes of the molecular crowding agents (or "crowders") affect IDPs is currently incomplete (26), especially for polymeric crowders. Here we use single-molecule spectroscopy to investigate the influence of crowding on the conformational distributions of IDPs, as a step toward a quantitative framework of how the polydisperse cellular environment affects these highly flexible molecules.

Single-molecule fluorescence detection in combination with Förster resonance energy transfer (FRET) is a method highly suited for addressing this question (5–7, 11, 27, 28) because it allows the heterogeneous structural ensemble of suitably labeled IDPs to be probed even in the presence of very large concentrations of unlabeled solutes. To investigate the physical principles underlying the crowding effects on IDPs, we study a selection of IDPs representative of the naturally occurring sequence compositions in combination with a broad range of molecular sizes of crowding agents. We primarily use polyethylene glycol (PEG) as a crowding agent. This uncharged polymer with high solubility in aqueous solution (29) (*SI Appendix*, Fig. S1) is available from monomeric ethylene glycol to degrees of polymerization of almost 1,000 (*SI Appendix*, Table S1) at sufficient purity for single-molecule experiments up to physiologically realistic volume fractions of crowder of ~40% (30). PEG is widely used for biomedical applications (31) and for mimicking inert crowding agents (13, 26). Previous work has shown that the conformational properties of IDPs strongly depend on their amino acid sequence composition and charge patterning (8, 11, 27, 28, 32–34). Here we investigate the effect of crowding on four different IDP sequences that span a broad range of net charge per residue and average hydrophobicity (Fig. 1 and *SI Appendix*, Table S2): the N- and C-terminal segments of human prothymosin-α (ProTα-N and -C), the binding domain of the activator for thyroid hormones and retinoid receptors (ACTR), and the N-terminal domain of HIV-1 integrase (IN). Whereas ProTα is highly charged and does not assume a folded structure under any known conditions, ACTR and IN are

Significance

In the interior of a cell, the volume accessible to each protein molecule is restricted by the presence of the large number of other macromolecules. Such a crowded environment is known to affect the stability and folding rates of proteins. In the case of intrinsically disordered proteins (IDPs), however, a class of proteins that lack stable structure, much less is known about the role of crowding effects. We have quantified the conformational changes occurring in IDPs in the presence of high concentrations of different polymers that act as crowding agents. Using single-molecule spectroscopy, we have identified effects that are typical of polymer solutions and have direct implications for the behavior of IDPs within the cell.

Author contributions: A.S. and B.S. designed research; A.S. and I.K. performed research; M.B.B., H.H., F.Z., and D.N. contributed new reagents/analytic tools; A.S. and I.K. analyzed data; and A.S., I.K., and B.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1] A.S. and I.K. contributed equally to this work.

[2] To whom correspondence may be addressed. E-mail: asoranno@bioc.uzh.ch or schuler@bioc.uzh.ch.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1322611111/-/DCSupplemental.

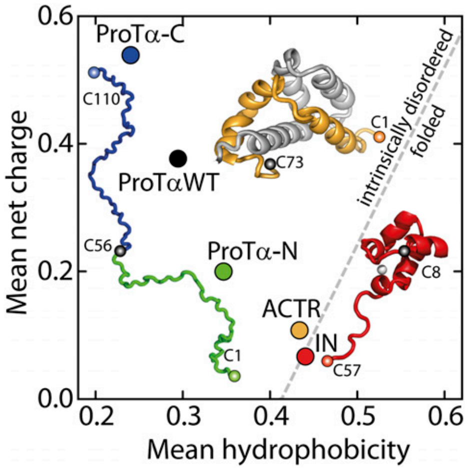


Fig. 1. Mean net charge versus mean hydrophobicity per residue for the four disordered protein sequences used in this study: the C- and N-terminal segments of prothymosin *α*, ProT*α*-C (blue) and ProT*α*-N (green), respectively (complete sequence: black), the activator for thyroid hormones and retinoid receptors, ACTR (orange), and the N-terminal domain of the HIV-1 integrase, IN (red). Folded structures refer to the conformations of ACTR and IN in presence of their ligands, NCBD (gray structure) and Zn^{2+} (light gray sphere). The FRET labeling sites (*SI Appendix*, Table S1) are indicated by colored spheres. The dashed gray line indicates the boundary between intrinsically disordered and folded proteins proposed by Uversky (32). Note that the contributions to the net charge from the fluorescent dyes are included (11).

representatives of the classes of IDPs that fold upon binding a protein or a small ligand, respectively.

Results

Quantifying Crowder-Induced Chain Compaction with Single-Molecule FRET. To probe the intramolecular distance distributions of the IDPs, we attached Alexa Fluor 488 and Alexa Fluor 594 as donor and acceptor fluorophores via cysteine residues introduced at suitable positions, with sequence separations of 55 (ProT*α*-N), 54 (ProT*α*-C), 72 (ACTR), and 49 residues (IN) (*SI Appendix*, Table S1). Fig. 2 shows examples of confocal single-molecule FRET experiments with the four different IDP sequences performed at increasing concentrations of PEG 6000 (i.e., PEG with a molecular mass of ∼6,000 Da; *SI Appendix*, Table S2). Up to three peaks are observed in the transfer efficiency (E) histograms from measurements of labeled IDPs freely diffusing in solution. The peak at $E \sim 0$ results from molecules lacking an active acceptor dye and is not of interest here. The second peak at intermediate E corresponds to the disordered state. The appearance of a third peak at $E \sim 0.7$ and $E \sim 0.9$ for ACTR and IN, respectively, results from the formation of a folded structure in complex with their ligands, the nuclear coactivator binding domain (NCBD) and a Zn^{2+} ion, respectively (*SI Appendix*). This separation of subpopulations is essential for distinguishing the effects of solutes on the conformational distributions within the disordered state from a cooperative transition to a folded state. In the case of IN, our experiments indicate the formation of a small folded population at high PEG concentrations even in the absence of Zn^{2+} (Fig. 2*D* and *SI Appendix*, Fig. S2), but for all other proteins, only an unfolded population is present (Fig. 2 and *SI Appendix*, Fig. S2). However, with increasing concentration of PEG 6000, three of the four disordered sequences (ProT*α*-C, ProT*α*-N, and ACTR) exhibit a clear shift of the peak corresponding to the disordered state toward higher transfer efficiencies, indicating an overall tendency of these proteins to collapse in the presence of crowding agents. For the case of IN, which has the least charged and most hydrophobic sequence (Fig. 1), only very small changes in transfer efficiency are noticeable, clearly demonstrating that molecular crowding does not affect all IDPs equally. Given the importance of intramolecular electrostatic repulsion for their conformations (11, 33), it may seem surprising that the more highly charged IDPs exhibit a more pronounced collapse.

The changes in transfer efficiency of the IDPs induced by the crowding agents can be used to extract information on the corresponding changes in chain conformations. Following previous work on unfolded proteins (35) and IDPs (27, 28), we use a Flory–Fisk distribution, which provides a description of the underlying distance distributions, to quantify the dimensions of the polypeptide chains in terms of mean-squared intramolecular distances or the effective radii of gyration, R_g, of the segments

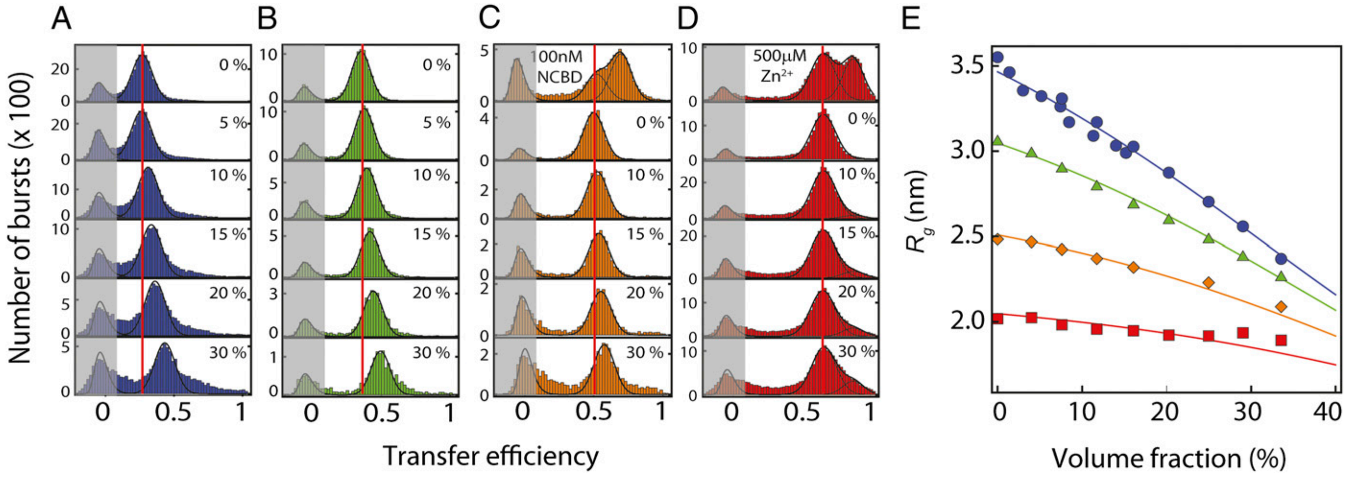


Fig. 2. Single-molecule FRET can be used for quantifying the compaction of disordered proteins by molecular crowding. Representative FRET efficiency histograms at different volume fractions of polyethylene glycol (PEG) 6000 for ProT*α*-C (*A*, blue), ProT*α*-N (*B*, green), ACTR (*C*, orange), and IN (*D*, red). Histograms of ACTR and IN in the presence of their respective interaction partners, NCBD (*C*) and Zn^{2+} (*D*), are shown for comparison. Gaussian and lognormal distributions are used to fit the data (solid lines). The transfer efficiency peaks from molecules lacking an active acceptor dye are shaded in gray. At the highest volume fractions of PEG, some broadening of the peaks is observed due to the increasing fluorescent background. Only IN exhibits a small crowder-induced population at the transfer efficiency of the folded state (see *SI Appendix*, Fig. S2 for detailed controls). (*E*) The resulting radii of gyration (R_g) for ProT*α*-C (blue circles), ProT*α*-N (green triangles), ACTR (orange rhombi), and IN (red squares) illustrate the PEG-induced compaction. Fits (solid lines) are obtained using scaled-particle theory (*SI Appendix*, Eq. **S5**) with the size of PEG 6000 as a single, globally adjustable fitting parameter. The precision of the values of R_g as estimated from multiple measurements of selected data points is comparable to or smaller than the size of the symbols.

probed by the FRET pair (*SI Appendix*). Note that the analysis is robust with respect to the polymer-physical model used and that the use of multiparameter detection allows us to exclude possible interfering artifacts, such as insufficient rotational averaging of the fluorophores or quenching of the dyes (*SI Appendix*).

Fig. 2*E* shows examples of the resulting changes in R_g as a function of the volume fraction ϕ of PEG 6000 for the four IDP sequences, all of which exhibit collapse upon crowding. Between 0% and 40% of crowder, the changes in R_g range from 0.2 nm (or ~10%) for IN to ~1 nm (or ~30%) for ProTα-C. Qualitatively, this is the behavior expected even from a simple hard-sphere model for a crowding agent whose steric repulsion of the IDP chains leads to their compaction (13, 36). A commonly used quantitative framework for such effects is scaled-particle theory (37), which provides an estimate of the change in free energy required for creating a cavity equivalent to the size of the IDP in a solution of hard spheres with a radius corresponding to the size of the crowding agent, R_g^{crd} (*SI Appendix*). If we apply scaled-particle theory, a remarkably good fit is achieved with R_g^{crd} as a global fit parameter (Fig. 2*E*). However, the resulting value for R_g^{crd} of (6.2 ± 0.1) nm is almost twice the measured radius of gyration of PEG 6000 (*SI Appendix*, Fig. S1), signifying that a hard-sphere description is not adequate for polymeric crowding agents such as PEG (38).

Crowder Size Variation Reveals the Importance of Polymer Effects. To identify the origin of this discrepancy, we choose a strategy orthogonal to varying the volume fractions of crowder and probe the influence of different sizes of crowding agents on the compaction of IDPs. Fig. 3 shows the complete data set for all four IDP sequences with PEGs of 10 different degrees of polymerization, P, at volume fractions from 0% to ~40%. For all IDPs, we observe the tendency to collapse with increasing crowder concentration, but interestingly, the degree of compaction is highly dependent on crowder size. The characteristic behavior is most apparent if we consider the change in R_g of an IDP as a function of P at a fixed volume fraction of PEG, as illustrated in Fig. 4 for ProTα-C with ϕ = 15%. The IDPs collapse monotonically as the crowder size increases, but their R_g reaches a plateau for PEGs of more than ~100 monomers. Notably, this behavior is the opposite of what we expect from scaled-particle theory because the free energy cost for creating a cavity of given size decreases with increasing crowder size (*SI Appendix*); in other words, larger solid-sphere crowding agents have larger interstitial cavities and would thus accommodate expanded IDPs more easily (Fig. 4*A*). To illustrate the discrepancy, Fig. 4*E* shows the resulting prediction for $R_g(P)$ based on scaled-particle theory (solid black line, Fig. 4*E*).

An obvious deficit of scaled-particle theory for the treatment of unfolded proteins is the assumption that the crowders cannot penetrate the unfolded chain. To address this issue, Minton proposed the "Gaussian cloud" model (37) (Fig. 4*B*), where the unfolded protein is described in terms of a continuous Gaussian distribution of monomer density around the center of mass of the protein (*SI Appendix*, Fig. S3). Small solid-sphere crowders can pervade this protein cloud and thus have little effect on the density distribution of the chain. With increasing crowder size, the probability of accommodating the corresponding spheres without steric clashes with the chain decreases, leading to a compaction of the IDP, in agreement with experimental observation (solid gray line, Fig. 4*E*). For very large crowding agents, however, this penetration probability decreases further, and ultimately the limit of classic scaled-particle theory is recovered, in contrast with the experimental observation.

These results strongly suggest that we need to go a step further and take into account the polymeric nature of both IDP and crowding agent to explain the behavior observed experimentally. The simplest realistic model needs to comprise two polymers of different lengths in good solvent, i.e., a ternary system. Note that both the IDPs (24) and the crowder (*SI Appendix*, Fig. S2) (26) exhibit the scaling behavior characteristic of polymers in good solvent, which justifies this assumption.

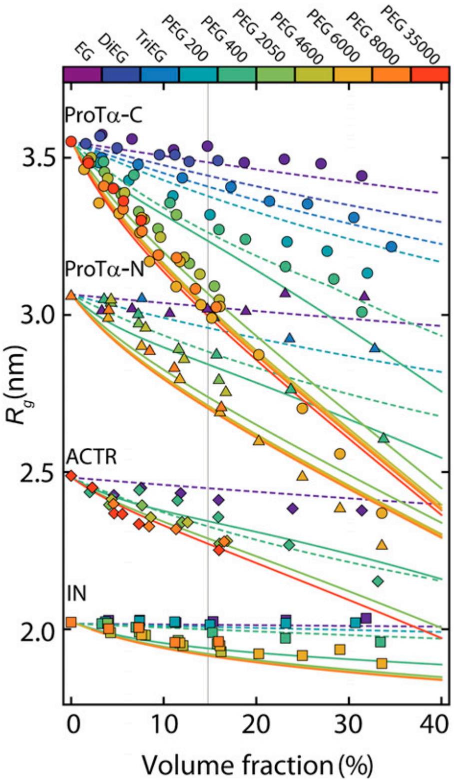


Fig. 3. Both increasing volume fraction and increasing crowder size lead to IDP compaction. Radii of gyration of ProTα-C (circles), ProTα-N (triangles), ACTR (rhombi), and IN (squares) as a function of the volume fraction of PEG obtained from single-molecule FRET experiments. Fits to the data corresponding to the short-chain regime (dashed lines, Eq. **1a**) and the long-chain regime (solid lines, Eq. **1b**) are shown. For the case of PEG 400, both types of fits are reported to illustrate the cross-over between the two regimes. The vertical dashed line indicates the volume fraction of 15% PEG used in Fig. 4.

We also need to take into consideration that, unlike the hard spheres assumed in scaled-particle theory, polymer chains can interpenetrate. This aspect becomes most relevant above a limiting volume fraction, referred to as the overlap concentration ϕ^*, where the solution can be thought of as being filled by nonintersecting spheres of the size of a single polymer chain. For volume fractions greater than ϕ^*, the transition between dilute and semidilute regimes occurs, and the chains start to overlap, which will affect the conformations of the polymers (*SI Appendix*, Fig. S1). ϕ^* depends only on the length P of the polymers and on the scaling exponent in the appropriate solvent regime ($\phi^* = P^{-4/5}$ in good solvent; *SI Appendix*); for long chains, this semidilute regime is reached already at volume fractions of a few percent (*SI Appendix*, Fig. S1) and the interpenetration of the chains must thus be taken into account for the majority of our experimental conditions.

Within the framework of the commonly used Flory–Huggins theories, we therefore need to distinguish two scenarios under our experimental conditions: the short-chain regime (Fig. 4*C*) and the long-chain regime (Fig. 4*D*) (39). In the first case, the crowding polymer chains are short and consequently remain below the overlap concentration. The system can thus be depicted as a dilute ($\phi < \phi^*$) solution of PEG chains of radius R_g^{crd} that do not overlap with each other but are able to pervade the volume explored by the IDP (Fig. 4*C*) (39). Inside this volume, the degrees of freedom of the crowders are reduced by the IDP, and the crowder chains will gain entropy by leaving this volume. A further increase in entropy of the crowder molecules results from reducing the volume occupied by the protein. In

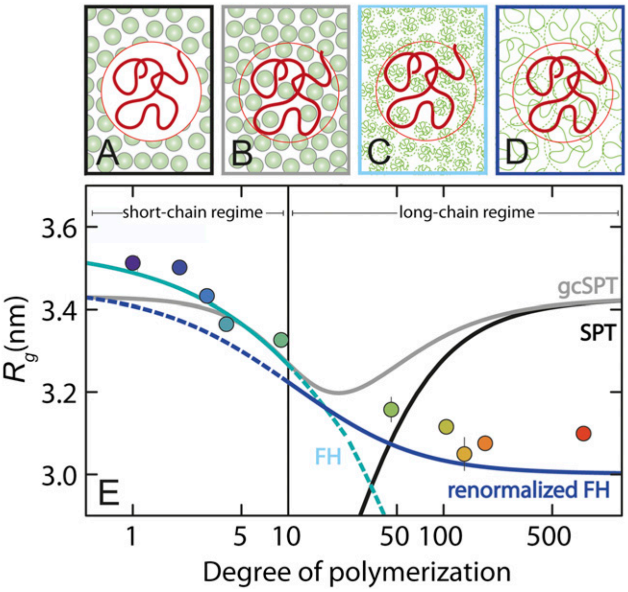


Fig. 4. Polymer concepts explain the compaction of IDPs by crowding agents of increasing size. Graphical representation of (*A*) scaled-particle theory (SPT), (*B*) Gaussian cloud model (gcSPT), (*C*) Flory–Huggins theory (FH) in the short-chain regime, and (*D*) renormalized Flory–Huggins theory (renormalized FH) in the long-chain regime. (*E*) Radius of gyration of ProTα-C as a function of the degree of polymerization of PEG at 15% volume fraction of crowding agent. The data points were obtained from linear interpolation of the volume fraction dependences shown in Fig. 3 (same color code for the PEG size). Fits according to the different theories are shown as black (SPT), gray (gcSPT), cyan (FH theory), and blue (renormalized FH theory) lines. Solid lines indicate the regime for which the respective theories were derived; outside of these regimes, dashed lines are used. Error bars reporting on the precision of the experiments are calculated as 1 SD from the linear fits of data for each PEG series in Fig. 3 (uncertainties are smaller than the size of the symbols unless shown explicitly).

other words, the requisite equality of chemical potentials for crowders inside and outside the volume pervaded by the IDP predicts a collapse of the protein chain (39), similar to the Gaussian cloud model, and in good agreement with the experimental data (cyan line, Fig. 4*E*; also see *SI Appendix*). In the long-chain regime, however, this mean-field theory fails and diverges from the measured results. In this regime, the crowding polymers are often above their overlap concentrations, and their conformations are influenced by mutual interpenetration. In contrast with the case of a single chain in good solvent, where the dimensions are dominated by repulsive interactions between the monomers, the interpenetration by other crowders in the semidilute regime causes a screening of these repulsive interactions within each chain (40, 41). This excluded volume screening will also affect the conformations of the IDP. However, because the polymers have dimensions comparable to or larger than the protein, they will only partially penetrate the IDP. Under these conditions, the ternary system is close to a critical point and can exhibit density fluctuations over a broad range of length scales due to interactions within the protein, within the crowders, and between the crowders and the protein (41, 42). Many critical systems, ranging from the liquid–gas phase transition near the critical point to the magnetization near the Curie point of a ferromagnet and the Kondo effect of electrons in metals, have been successfully described by renormalization group theory (43). The same approach has provided fundamental insights into the scaling invariance for polymer solutions (41). Here we adopt a renormalized Flory–Huggins-type theory developed by Schäfer and Kappeler (44) for a multicomponent system in the long-chain regime.

We thus analyzed the data in the short-chain and long-chain regimes according to

$$R_g(N,P,\phi,a)=R_{g0}\left(\frac{1}{1+a\phi/\phi^*(P)}\right)^{1/5}\quad \text{for}\quad P<N^{1/2},\qquad \textbf{[1a]}$$

and

$$R_g(N,P,\phi,s_{NP})=R_{g0}f(N,P,\phi,s_{NP})\quad \text{for}\quad P\geq N^{1/2},\qquad \textbf{[1b]}$$

where R_{g0} is the radius of gyration of the IDP in the absence of crowding; a is an empirical parameter that can account for differences in the solvent quality for the different proteins and interactions between protein and polymer (45) (*SI Appendix*); s_{NP} quantifies the interaction between the protein and the polymer chains; and f is a function that represents the renormalization mapping (*SI Appendix*). It is worth emphasizing that Eqs. **1a** and **1b** contain only a single adjustable parameter each, a and s_{NP}, respectively (*SI Appendix*, Table S4 and Fig. S4). The equations provide a good fit to the experimental data, including the approach to a limiting value of R_g for IDPs in very large crowders (Fig. 4*E*). In fact, the entire data set for all four IDP sequences is described remarkably well by a global fit (Fig. 3). The success of this approach supports the hypothesis that the polymeric properties of both IDP and PEG are essential for understanding the effect of molecular crowding, and that the criticality of the solution cannot be neglected. Considering the highly polydisperse cellular environment, it seems probable that related effects will be prominent in vivo and that mean-field descriptions are insufficient for a quantitative description of crowding in the cell.

The Balance of Hard-Core Repulsion and Other Nonspecific Interactions. Recent experimental results indicate that the presence of weak, nonspecific attractive interactions in the heterogeneous cellular environment can modulate or even dominate the effects of hard-core repulsion that are at the basis of molecular crowding (46–48). The role of such "chemical interactions" is a subject of debate also for proteins and PEG (13, 26). Notably, the approach presented here (Eq. **1b**) allows the relative contributions of hard-core repulsion and other interactions to be quantified in terms of the interaction parameter s_{NP}. In the cases investigated here, the analysis with Eq. **1b** indicates that a small contribution of unfavorable interactions with PEG is present for ProTα and ACTR, and no such interactions are detected in the case of IN (*SI Appendix*, Table S4). We note, however, that even though interactions such as nonspecific attraction between crowder and IDP can modulate the amplitude of the change in R_g with crowder concentration (*SI Appendix*, Fig. S4), the polymeric effects dominate the overall behavior.

An independent means of interrogating the role of nonspecific charge and hydrophobic interactions is to add salt or denaturants to the solution. Fig. 5 shows that neither 1 M KCl nor 4 M guanidinium chloride (GdmCl) nor 4 M urea impedes the collapse of ProTα. The value of R_{g0} depends on ionic strength and denaturant concentration owing to the known charge screening and/or denaturant-induced chain expansion (11). However, the dependence of R_g on the volume fraction of PEG is described by Eq. **1b** with the same values of s_{NP} as in the absence of salt or denaturant, just by rescaling R_{g0} to the value at the corresponding KCl, GdmCl, or urea concentrations without crowder, suggesting that the effect of additional interactions on the compaction of the IDP is small. Finally, we tested the influence of different chemical structures of the crowding polymer in experiments with dextran, polyvinyl alcohol (PVA), and polyvinylpyrrolidone (PVP) (Fig. 5). Even though we could measure these solutions only for volume fractions of crowder of up to 10% owing to fluorescent impurities, in all cases we observed a collapse of ProTα-C similar to that in PEG. The resulting values of s_{NP} for dextran, PVA, and PVP are significantly lower than for PEG (*SI Appendix*, Table S5), indicating better compatibility—or less unfavorable interactions—with ProTα, but the collapse of the IDP is preserved. In summary, the polymeric crowding effects on IDPs observed here are dominated

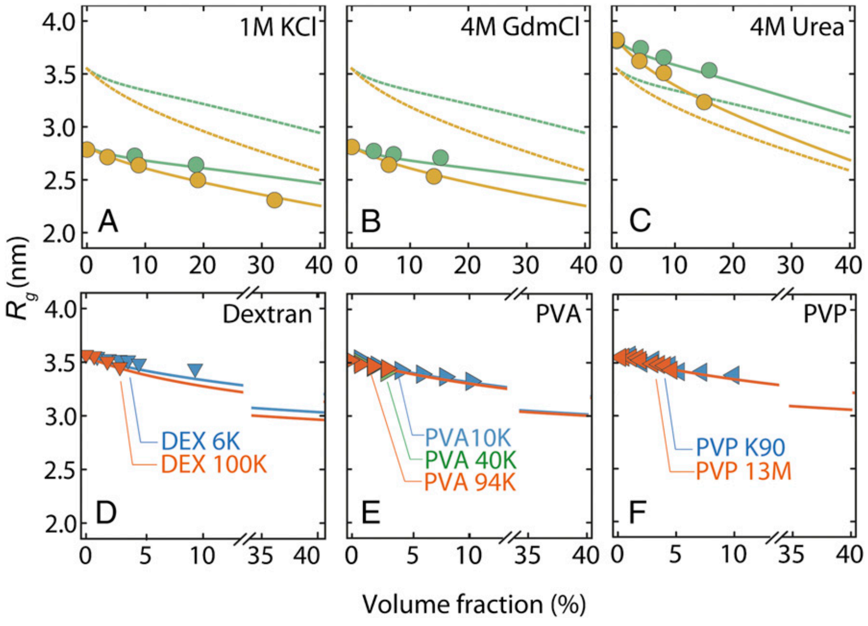


Fig. 5. Variation of solution conditions and crowding agents suggest the importance of nonspecific effects on IDP compaction. Radius of gyration of ProTα-C versus the volume fraction of PEG 400 (green circles) and PEG 6000 (yellow circles) in (*A*) 1 M KCl solution, (*B*) 4 M GdmCl, and (*C*) 4 M urea. Fits according to Eq. **1b**, assuming a different value of R_{g0} but the same value of s_{NP} as in Figs. 3 and 4, are shown as green and yellow solid lines for PEG 400 and PEG 6000, respectively. Fits for the same crowding agents in the absence of salt or denaturant (Fig. 3) are included as dashed lines with corresponding colors. The effects of molecular crowding with dextran (*D*), PVA (*E*), and PVP (*F*) on ProTα-C are shown for different sizes of these alternative crowders as indicated. Lines represent the fit to the renormalized FH theory (*SI Appendix*, Table S5) and are extrapolated up to 40% volume fraction for comparison with other polymers.

by hard-core repulsion between the monomers and the resulting excluded volume screening (40, 41), indicating a phenomenon of generic relevance. However, the analysis presented here does allow additional interactions to be included that can modulate the crowding effect.

Discussion

Eqs. **1a** and **1b** can account for the dependence of R_g on crowder concentration and crowder size for all four IDPs investigated (Fig. 3). The question remains, however, why the extent of crowder-induced compaction is so different for the different IDPs. Polymer theory offers an interesting explanation. According to the Flory theorem, the chains in a melt (i.e., in the absence of solvent) of compatible polymers approach their Θ-state. Under these conditions, because of the screening of excluded volume interactions within and between the polymers, the dimensions of the chains scale approximately with the square root of the number of chain segments, and a characteristic radius of gyration $R_{g\Theta}$ is observed (*SI Appendix*). Recent work indicates that $R_{g\Theta}$ for the IDPs investigated here is in the range of ~1.7–2.0 nm (27) (*SI Appendix*). The results in Fig. 3 for the larger PEGs are indeed consistent with asymptotic convergence of R_g for all of the IDP variants toward values in this range in the limit of very high volume fractions of crowder, i.e., under conditions that approach the situation of a melt. In other words, highly expanded IDPs with dimensions much greater than $R_{g\Theta}$ (such as ProTα) are expected to undergo more pronounced compaction on polymeric crowding than those IDPs that are close to $R_{g\Theta}$ already in the absence of crowders (such as IN). Based on the empirical relations between solvent quality and average net charge obtained previously (27), we estimate that ~90% of all IDPs are above the θ-state in the absence of crowding (*SI Appendix*) and should thus be susceptible to compaction by polymeric crowders.

The observations reported here could thus have implications for the functional properties of many IDPs, e.g., for the capture radius for their cellular targets in the framework of a fly-casting mechanism (49, 50) and for the folding propensity of denatured ensembles in the crowded cellular environment (13). However, the balance of the different contributions may be subtle. Whereas a compaction of the chain by crowding will result in a decrease of the capture radius, it will increase the translational diffusion coefficient. These opposing effects will modulate the basic influence of crowding on solution viscosity and the concomitant changes in association rates (51). Similarly, the established effects of crowding on the stability of the folded and/or bound states of IDPs (13) may be affected by changes in unfolded state dimensions. Single-molecule experiments of the type presented here may help to dissect these contributions quantitatively. Complementary simulations of polymeric crowding could provide valuable insights into the underlying molecular mechanisms.

We note that a substantial fraction of crowding in the cell is due to polymeric molecules such as peptides, nucleic acids, polysaccharides, or other disordered proteins. However, the extent of crowding is strongly affected by the spatial organization of the cell. A remarkable example of very high local concentrations of IDPs are nucleoporins, which line the nuclear pore complexes (25). We estimate the volume fraction occupied by nucleoporins to be between 25% and 55% of the volume available in the pore, about an order of magnitude greater than the overlap concentration (*SI Appendix*). Similarly, IDPs involved in RNA granules (22, 24) or analogous nonmembrane-bound bodies with liquid-like properties (23, 24) are likely to exceed their overlap concentration locally (*SI Appendix*). Under these conditions, polymer effects characteristic of the semidilute regime will be highly relevant for the conformations of IDPs and for the occurrence of possible phase transitions. Interestingly, ProTα often colocalizes with dense speckles such as promyelocytic leukemia bodies (52). Given its abundance in the nucleus of mammalian cells and its high mobility within and near the nucleus (53), we expect that a compaction similar to what we observed here can occur in vivo. According to our results, the dense local environment resulting from liquid–liquid demixing (23, 24) or sol–gel transitions (22) should strongly influence the conformational distributions of IDPs, with consequent impact on the functional properties of the resulting assemblies and their mechanisms of formation. Flory–Huggins theories as used here might thus provide novel insights into the demixing of multicomponent polymeric systems (41). An interesting next step will be a direct comparison of experiments in vitro with intracellular measurements (14, 26), and the required quantitative tools are beginning to emerge (54–56).

Methods

Proteins were expressed, purified, and labeled similar to previous reports (11, 27, 28). Single-molecule measurements were performed using a MicroTime 200 confocal microscope equipped with a HydraHarp 400 counting module (PicoQuant). For details on experiments and theory, see *SI Appendix*.

ACKNOWLEDGMENTS. We thank Rohit Pappu, Devarajan Thirumalai, and David Goldenberg for helpful discussions and comments on the manuscript. This work was supported by the Swiss National Science Foundation and a Starting Investigator Grant of the European Research Council.

1. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6(3):197–208.
2. Dunker AK, Silman I, Uversky VN, Sussman JL (2008) Function and structure of inherently disordered proteins. *Curr Opin Struct Biol* 18(6):756–764.
3. Jensen MR, Ruigrok RW, Blackledge M (2013) Describing intrinsically disordered proteins at atomic resolution by NMR. *Curr Opin Struct Biol* 23(3):426–435.
4. Wright PE, Dyson HJ (2009) Linking folding and binding. *Curr Opin Struct Biol* 19(1): 31–38.
5. Ferreon AC, Moran CR, Gambin Y, Deniz AA (2010) Single-molecule fluorescence studies of intrinsically disordered proteins. *Methods Enzymol* 472:179–204.
6. Schuler B, Müller-Späth S, Soranno A, Nettels D (2012) Application of confocal single-molecule FRET to intrinsically disordered proteins. *Methods Mol Biol* 896:21–45.
7. Ferreon AC, Ferreon JC, Wright PE, Deniz AA (2013) Modulation of allostery by protein intrinsic disorder. *Nature* 498(7454):390–394.
8. Mao AH, Lyle N, Pappu RV (2013) Describing sequence-ensemble relationships for intrinsically disordered proteins. *Biochem J* 449(2):307–318.
9. Lyle N, Das RK, Pappu RV (2013) A quantitative measure for protein conformational heterogeneity. *J Chem Phys* 139(12):121907.
10. Fisher CK, Stultz CM (2011) Protein structure along the order-disorder continuum. *J Am Chem Soc* 133(26):10022–10025.
11. Müller-Späth S, et al. (2010) From the Cover: Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc Natl Acad Sci USA* 107(33): 14609–14614.
12. Uversky VN (2009) Intrinsically disordered proteins and their environment: Effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein J* 28(7-8):305–325.
13. Zhou HX, Rivas GN, Minton AP (2008) Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *Annu Rev Biophys* 37:375–397.
14. Gershenson A, Gierasch LM (2011) Protein folding in the cell: Challenges and progress. *Curr Opin Struct Biol* 21(1):32–41.
15. Dedmon MM, Patel CN, Young GB, Pielak GJ (2002) FlgM gains structure in living cells. *Proc Natl Acad Sci USA* 99(20):12681–12684.
16. McNulty BC, Young GB, Pielak GJ (2006) Macromolecular crowding in the Escherichia coli periplasm maintains alpha-synuclein disorder. *J Mol Biol* 355(5):893–897.
17. Munishkina LA, Cooper EM, Uversky VN, Fink AL (2004) The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *J Mol Recognit* 17(5): 456–464.
18. Szasz CS, et al. (2011) Protein disorder prevails under crowded conditions. *Biochemistry* 50(26):5834–5844.
19. Hong JA, Gierasch LM (2010) Macromolecular crowding remodels the energy landscape of a protein by favoring a more compact unfolded state. *J Am Chem Soc* 132(30):10445–10452.
20. Mikaelsson T, Adén J, Johansson LBA, Wittung-Stafshede P (2013) Direct observation of protein unfolded state compaction in the presence of macromolecular crowding. *Biophys J* 104(3):694–704.
21. Johansen D, Jeffries CM, Hammouda B, Trewhella J, Goldenberg DP (2011) Effects of macromolecular crowding on an intrinsically disordered protein characterized by small-angle neutron scattering with contrast matching. *Biophys J* 100(4):1120–1128.
22. Han TW, et al. (2012) Cell-free formation of RNA granules: Bound RNAs identify features and components of cellular assemblies. *Cell* 149(4):768–779.
23. Li PL, et al. (2012) Phase transitions in the assembly of multivalent signalling proteins. *Nature* 483(7389):336–340.
24. Brangwynne CP (2011) Soft active aggregates: Mechanics, dynamics and self-assembly of liquid-like intracellular protein bodies. *Soft Matter* 7(7):3052–3059.
25. Rout MP, et al. (2000) The yeast nuclear pore complex: Composition, architecture, and transport mechanism. *J Cell Biol* 148(4):635–651.
26. Elcock AH (2010) Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. *Curr Opin Struct Biol* 20(2):196–206.
27. Hofmann H, et al. (2012) Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proc Natl Acad Sci USA* 109(40): 16155–16160.
28. Soranno A, et al. (2012) Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy. *Proc Natl Acad Sci USA* 109(44): 17800–17806.
29. Devanand K, Selser JC (1991) Asymptotic-behavior and long-range interactions in aqueous-solutions of poly(ethylene oxide). *Macromolecules* 24(22):5943–5947.
30. Zimmerman SB, Trach SO (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. *J Mol Biol* 222(3):599–620.
31. Harris JM (1992) *Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications* (Plenum, New York).
32. Uversky VN, Gillespie JR, Fink AL (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* 41(3):415–427.
33. Mao AH, Crick SL, Vitalis A, Chicoine CL, Pappu RV (2010) Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc Natl Acad Sci USA* 107(18):8183–8188.
34. Das RK, Pappu RV (2013) Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc Natl Acad Sci USA* 110(33):13392–13397.
35. Ziv G, Haran G (2009) Protein folding, protein collapse, and Tanford's transfer model: Lessons from single-molecule FRET. *J Am Chem Soc* 131(8):2942–2947.
36. Cheung MS, Klimov D, Thirumalai D (2005) Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc Natl Acad Sci USA* 102(13): 4753–4758.
37. Minton AP (2005) Models for excluded volume interaction between an unfolded protein and rigid macromolecular cosolutes: Macromolecular crowding and protein stability revisited. *Biophys J* 88(2):971–985.
38. Mittal J, Best RB (2010) Dependence of protein folding stability and dynamics on the density and composition of macromolecular crowders. *Biophys J* 98(2):315–320.
39. Joanny JF, Grant P, Pincus P, Turkevich LA (1981) Conformations of polydisperse polymer-solutions - bimodal distribution. *J Appl Phys* 52(10):5943–5948.
40. Edwards SF (1966) Theory of polymer solutions at intermediate concentration. *Proc Phys Soc Lond* 88(560P):265–280.
41. Schäfer L (1999) *Excluded Volume Effects in Polymer Solutions as Explained by the Renormalization Group* (Springer, Berlin).
42. Tran HT, Pappu RV (2006) Toward an accurate theoretical framework for describing ensembles for proteins under strongly denaturing conditions. *Biophys J* 91(5):1868–1886.
43. Wilson KG (1983) The renormalization-group and critical phenomena. *Rev Mod Phys* 55(3):583–600.
44. Schäfer L, Kappeler C (1993) Interaction effects on the size of a polymer-chain in ternary solutions - a renormalization-group study. *J Chem Phys* 99(8):6135–6154.
45. Nose T (1986) Chain dimension of a guest polymer in the semidilute solution of compatible and incompatible polymers. *J Phys (Paris)* 47(3):517–527.
46. Minton AP (2013) Quantitative assessment of the relative contributions of steric repulsion and chemical interactions to macromolecular crowding. *Biopolymers* 99(4): 239–244.
47. Sarkar M, Li C, Pielak GJ (2013) Soft interactions and crowding. *Biophys. Rev.* 5(2): 187–194.
48. Kim YC, Mittal J (2013) Crowding induced entropy-enthalpy compensation in protein association equilibria. *Phys Rev Lett* 110(20):208102-1–208102-5.
49. Shoemaker BA, Portman JJ, Wolynes PG (2000) Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc Natl Acad Sci USA* 97(16): 8868–8873.
50. Trizac E, Levy Y, Wolynes PG (2010) Capillarity theory for the fly-casting mechanism. *Proc Natl Acad Sci USA* 107(7):2746–2750.
51. Schreiber G, Haran G, Zhou HX (2009) Fundamental aspects of protein-protein association kinetics. *Chem Rev* 109(3):839–860.
52. Vareli K, Frangou-Lazaridis M, van der Kraan I, Tsolas O, van Driel R (2000) Nuclear distribution of prothymosin alpha and parathymosin: Evidence that prothymosin alpha is associated with RNA synthesis processing and parathymosin with early DNA replication. *Exp Cell Res* 257(1):152–161.
53. Enkemann SA, Ward RD, Berger SL (2000) Mobility within the nucleus and neighboring cytosol is a key feature of prothymosin-alpha. *J Histochem Cytochem* 48(10): 1341–1355.
54. Gelman H, Platkov M, Gruebele M (2012) Rapid perturbation of free-energy landscapes: From in vitro to in vivo. *Chemistry* 18(21):6420–6427.
55. Phillip Y, Kiss V, Schreiber G (2012) Protein-binding dynamics imaged in a living cell. *Proc Natl Acad Sci USA* 109(5):1461–1466.
56. Sakon JJ, Weninger KR (2010) Detecting the conformation of individual proteins in live cells. *Nat Methods* 7(3):203–205.

Supplementary Information

Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments

Andrea Soranno[1], Iwo König[1], Madeleine B. Borgia, Hagen Hofmann, Franziska Zosel, Daniel Nettels, and Benjamin Schuler

[1] Both authors contributed equally to this work

(Proceedings of the National Academy of Sciences of the USA, 2014)

Supporting Information

Material and Methods

Protein preparation and labeling. Cysteine residues for the specific labeling of IDPs using maleimide chemistry were introduced by site-directed mutagenesis at the positions given in Table S1. ProTα (57) variants, IN (58) and ACTR (59) were produced in E. coli BL21 with an N-terminal His-tag for purification. Cells were grown in LB medium and expression was induced with IPTG.

Disruption of harvested cells and ammonium sulfate precipitation were carried out as described previously for ProTα (60). The protein was bound to a gravity flow Ni-NTA column (Thermo Scientific) and eluted with 20 mM Tris, 100 mM sodium chloride, and 500 mM imidazole, pH 7.5. The sample was dialyzed against 20 mM Tris, 100 mM sodium chloride, and 20 mM imidazole, pH 7.5. The His-tag was cleaved off with HRV 3C protease, which also contained a His-tag; a second Ni-NTA chromatography run was used to remove the protease and the His-tag. The flow-through containing ProTα was concentrated with a YM-3 Centriprep centrifugal filter (Merck Millipore), reduced with 10 mM TCEP, and purified by reversed phase (RP) HPLC on a Reprosil Gold 200 column (Dr. Maisch, Germany) and elution with an acetonitrile gradient. Purified ProTα was lyophilized in a SpeedVac concentrator (Thermo Scientific), dissolved in 0.1 M sodium phosphate, pH 7.5, and labeled with Alexa Fluor 488 maleimide (Invitrogen) at a molar ratio of dye to protein of 0.8:1. The reaction mixture was quenched with β-mercaptoethanol, reduced with 10 mM TCEP and purified by RP-HPLC on an XTerra C18 column (Waters). The fraction containing singly labeled ProTα was lyophilized in a SpeedVac concentrator, and labeled and purified analogously with a molar excess of Alexa Fluor 594 maleimide (Invitrogen). The masses of the doubly labeled ProTα-N and ProTα-C were confirmed by electrospray ionization mass spectrometry (ESI-MS).

IN was purified by Ni-NTA chromatography as described for ProTα. The His-tag was removed by enzymatic cleavage with Thrombin protease after dialysis against 50 mM $NaHCO_3$ pH 9.3, 0.5 M Na_2SO_4, 0.1 M NaCl, 5 mM EDTA, 1 mM DTT. After reduction with 10 mM TCEP and purification by RP-HPLC, IN was lyophilized in a SpeedVac concentrator, dissolved in 50 mM HEPES, 0.5 mM $ZnCl_2$, and 0.5 M arginine, pH 7.5, and labeled with Alexa Fluor 488 maleimide at a molar ratio of dye to protein of 0.8:1. The reaction mixture was quenched with β-mercaptoethanol, reduced with 10 mM TCEP and purified by RP-HPLC on a XTerra C18 column. The fraction containing singly labeled IN was lyophilized in a SpeedVac concentrator, and labeled and purified analogously with Alexa Fluor 594 maleimide. The doubly labeled IN was purified by RP-HPLC on an XTerra C18 column. The correct mass of the labeled IN was confirmed by ESI-MS.

ACTR was co-expressed with NCBD (nuclear co-activator binding domain of CREB) to improve the stability of ACTR during expression (59). The harvested cells were disrupted with a TS 1.1 cell disruption system (Constant Systems Ltd, England), the protein was bound to a Ni-NTA column, and the His-tag was removed by enzymatic cleavage with HRV 3C protease. After reduction of the sample with β-mercaptoethanol, the protease and the His-tag were removed with a second Ni-NTA column. ACTR was separated from NCBD by RP-HPLC on a Reprosil Gold 200 column by elution with an acetonitrile gradient. The fraction containing ACTR was lyophilized in a SpeedVac concentrator, dissolved in 0.1 M sodium phophate, pH 7.5, and labeled with Alexa Fluor 488 maleimide at a molar ratio of dye to protein of 0.8:1. The reaction mixture was again reduced with β-mercaptoethanol and purified by RP-HPLC on a Reprosil Gold 200 column. The fraction with the singly labeled ACTR was lyophilized in a SpeedVac concentrator, dissolved in labeling buffer and labeled with a molar excess of Alexa Fluor 594 maleimide. Doubly labeled ACTR was purified by RP-HPLC on a Reprosil Gold 200 column, and the correct mass was confirmed by ESI-MS.

Analogously, the same IDP variants were labeled with a different FRET pair (ATTO 532 and ATTO 647N, Atto-Tec, Germany) whose spectra are shifted towards higher wavelengths where the influence of fluorescent impurities on the transfer efficiency histograms at high PEG concentrations is strongly reduced.

Preparation of crowding solutions. Crowding experiments were carried out in 50 mM sodium phosphate buffer, pH 7.0. Crowding solutions were prepared by mixing acidic (50 mM NaH_2PO_4 + crowding agent) and alkaline (50 mM Na_2HPO_4 + crowding agent) stock solutions to a final pH of 7.0 (+/- 0.2). The concentration of the stock solutions depended on the solubility and purity of the crowding agent used (between 10 % w/w and 40 % w/w crowding agent). Crowding solutions with lower concentrations were prepared by dilution of the corresponding stock solution to the desired concentration with 50 mM sodium phosphate buffer, pH 7.0.

Single-molecule fluorescence spectroscopy. Single-molecule fluorescence measurements were performed with a MicroTime 200 confocal microscope (PicoQuant, Germany) equipped with a diode laser (LDH-D-C-485, PicoQuant, Germany), a 20 MHz supercontinuum laser (SC-450-4, Fianium, UK; wavelength selected with a z582/15 band pass filter (Chroma)) for pulsed interleaved excitation (PIE) (61), and an Olympus UplanApo 60x/1.20W objective (Olympus). Photons emitted from the sample were collected by the same objective. Remaining excitation light was eliminated by a filter (HQ500LP, Chroma Technology) before the emitted photons passed the confocal unit with a 100 µm pinhole. The emitted photons were separated into four channels with a polarizing beam splitter and a dichroic mirror (585DCXR, Chroma). Donor photons were filtered (ET525/50m, Chroma Technology) and then focused on a τ-SPAD avalanche photodiode (PicoQuant). Acceptor photons were filtered (HQ650/100m, Chroma Technology) and detected by a τ-SPAD (PicoQuant). The arrival time of every detected photon was recorded with a HydraHarp 400 counting module (PicoQuant).

All measurements were performed by exciting the donor dye with a laser power of 100 µW at the back aperture of the objective. For PIE measurements, the power used for exciting the acceptor dye was adjusted to match the intensity of the donor emission (between 50 and 70 µW). Single-molecule FRET efficiency histograms were acquired in samples with protein concentrations of about 50 pM to 100 pM. The time between excitation pulse and photon detection was stored with 16 ps resolution, with the lasers pulsed at a repetition rate of 20 MHz. The measurements were performed in 50 mM sodium phosphate buffer, pH 7.0, 200 mM β-mercaptoethanol, and 0.001% Tween-20 with varying concentrations of crowding agents and/or denaturants (guanidinium chloride, urea) or potassium chloride. Each sample was measured for 30 min to 1h at 295 K.

Data Analysis

FRET efficiency histograms. Fluorescence bursts from individual molecules were identified by combining successive photons separated by inter photon times of <100 µs and retaining the burst if the total number of photons detected after donor excitation was >50. Transfer efficiencies for each burst were calculated according to $E=n_A/(n_A+n_D)$, where n_D and n_A are the numbers of donor and acceptor photons, respectively. Corrections for background, acceptor direct excitation, channel crosstalk, differences in detector efficiencies, and quantum yields of the dyes were applied (6). The precision of the measurements as estimated from multiple independent measurements is typically ±0.01 transfer efficiency units and thus comparable to or smaller than the data points reported in the figures, unless shown explicitly. We estimate a systematic error for R_g of less than ±0.2 nm for the entire dataset. However, a uniform shift of this magnitude for all data toward higher or lower values of R_g does not affect any of our conclusions.

The changes in refractive index caused by increasing concentrations of crowding agents were measured with a digital Abbe refractometer (Krüss, Germany) and were used to recalculate the Förster radius (R_0) for every sample under the assumption that the polymeric crowding agents pervade the solutions uniformly. This assumption does not affect our conclusions, since the narrow range of refractive indices between 1.34 and 1.39 for our experimental conditions has a minor effect on the dimensions of the proteins. Even if we assumed the extreme case that no refractive index change was experienced by the protein locally upon going from pure buffer to the highest volume fraction of PEG, the lack of a refractive index correction would correspond to an apparent compaction of the protein between 0.05 nm (at transfer efficiency of 0.8) and 0.1 nm (at transfer efficiency of 0.3) relative to the values reported here.

Fluorescence lifetimes and anisotropy. Multiparameter detection allows us to exclude possible interfering artifacts, such as insufficient rotational averaging of the fluorophores or quenching of the dyes (62). The dependence of the fluorescence lifetimes on transfer efficiencies determined for each burst (Fig. S5) was compared with the behavior expected for fixed distances and for a chain sampling a broad distribution of distances. For a fixed distance, r, the mean donor lifetime in the presence of acceptor is given by $\tau_{DA}(r) = \tau_D\,(1-E(r))$, where τ_D is the lifetime in the absence of acceptor, and $E(r) = 1/(1+r^6/R_0^6)$. For a chain with a dye-to-dye distance distribution $P(r)$, the donor lifetime is $\tau_{DA} = \int_0^\infty tI(t)dt \,/ \int_0^\infty I(t)dt$, where $I(t) = I_0 \int_0^\infty P(r)e^{-t/\tau_{DA}(r)}dr$ is the time-resolved fluorescence emission intensity following donor excitation. Donor and acceptor lifetimes at different concentrations of crowding agents were analyzed by fitting subpopulation-specific time-correlated photon counting histograms after donor and acceptor excitation, respectively. This allows us to also examine the dependence of donor and acceptor lifetimes on the solution conditions. A systematic decrease of both donor and acceptor lifetimes of up to 10 % was observed with increasing concentrations of crowding agents. These lifetime changes are consistent with the changes expected according to the Strickler-Berg equation (63) for solutions with different refractive indices. Since the changes in donor and acceptor lifetimes are very similar, the contribution of this effect to the observed transfer efficiencies cancels. The variation of the donor lifetime also has no significant impact on R_0, since the donor quantum yield enters into the calculation of R_0 with the power of 1/6, resulting in a maximum change in R_0 of about 2%, less than the statistical experimental uncertainty.

Subpopulation-specific anisotropies were determined for both donor and acceptor, and values were found to vary between 0.03 and 0.08 for the donor and between 0.11 and 0.18 for the acceptor, consistent with values observed in ensemble measurements, and sufficiently low to assume as a good approximation for the orientational factor $\kappa^2 = 2/3$.

Quantifying the radius of gyration from transfer efficiencies. Essentially as described previously (27), FRET efficiencies are converted to radii of gyration according to

$$E = \int_0^{l_c} E(r) \int_{R_c}^{l_c/2} P(r \mid R_g) P_{FF}(R_g)\, dR_g dr, \qquad \text{(Eq. S1)}$$

where l_c is the contour length, and R_c is the radius of the sphere with volume equivalent to the sum of the volumes V_{aa}^i of all the amino acids, $R_c = \left(\frac{3}{4\pi} \sum_i V_{aa}^i \right)^{1/3}$. We use the conditional probability density function for a certain end-to-end distance, r, given the radius of gyration, R_g, suggested by Ziv and Haran (35), which describes the distance distribution of two random points inside the sphere of radius $\sqrt{5} \cdot r_g$ ($\sqrt{5}$ is a scaling factor used to satisfy the condition $6\left\langle R_g^2 \right\rangle = \left\langle r^2 \right\rangle$),

$$P\left(r \mid R_g\right) = \frac{1}{\sqrt{5}R_g} \left[3\left(\frac{r}{\sqrt{5}R_g} \right)^2 - \frac{9}{4}\left(\frac{r}{\sqrt{5}R_g} \right)^3 + \frac{3}{16}\left(\frac{r}{\sqrt{5}R_g} \right)^5 \right], 0 \le r < 2\sqrt{5}R_g \quad \text{(Eq. S2)}$$

$P_{FF}(R_g)$ is the Flory-Fisk distribution for the radius of gyration

$$P_{FF}(R_g) = Z^{-1} R_g^6 e^{-\frac{7R_g^2}{2\left\langle R_g^2 \right\rangle}}, \quad \text{(Eq. S3)}$$

where $\left\langle R_g^2 \right\rangle^{1/2}$ is the root-mean-squared radius of gyration of the chain, and Z is the normalization term. In contrast to previous works, we do not apply corrections to the Flory-Fisk distribution introduced by Sanchez theory (35, 64, 65), since the applied weighting proposed by Sanchez would be valid only for the measurements in absence of crowders and would be inconsistent with the other experimental conditions considered here (see *Scaled-particle theory* section). However, using the Sanchez distribution would result only in a systematic shift of all radii of gyration by approximately 0.1-0.2 nm, which does not affect any conclusions of this work.

Scaled-particle theory (SPT). Following the approach proposed by Minton (37), the effect of macromolecular crowding on the unfolded state can be quantified by weighting the probability density function of the radius of gyration of the disordered ensemble, $P(R_g)$, according to the chemical potential, $\Delta\mu = \Delta\mu(\phi, R_g, R_g^{crd})$, obtained with SPT:

$$\left\langle R_g^2 \right\rangle = \frac{\int R_g^2 P(R_g) e^{-\frac{\Delta\mu(\phi, R_g, R_g^{crd})}{k_B T}} dR_g}{\int P(R_g) e^{-\frac{\Delta\mu(\phi, R_g, R_g^{crd})}{k_B T}} dR_g}. \quad \text{(Eq. S4)}$$

The simplest choice for $P(R_g)$ is the Flory-Fisk distribution (Eq. S3), where $\left\langle R_g^2 \right\rangle^{1/2}$ is substituted by the root-mean-squared radius of gyration of the chain in absence of crowding agents, $\left\langle R_{g0}^2 \right\rangle^{1/2}$. If we assume that both the unfolded protein and the crowding agent can be described as rigid spheres (as in the classic SPT), the excess chemical potential of the IDP in the presence of crowding can be written as:

$$\frac{\Delta\mu(\phi, R_g, R_g^{crd})}{k_B T} = -\log\left(1-\phi\right) + \left(\overline{R}^3 + 3\overline{R}^2 + 3\overline{R}\right)\frac{\phi}{1-\phi} + \left(2\overline{R}^3 + \frac{9}{2}\overline{R}^2\right)\left(\frac{\phi}{1-\phi}\right)^2 + 3\overline{R}^3\left(\frac{\phi}{1-\phi}\right)^3, \text{(Eq. S5)}$$

where $\overline{R} = \frac{R_g}{R_g^{crd}}$; ϕ is the volume fraction of crowding agent in solution; and R_g^{crd} the radius of gyration of the crowder. Consequently, an increase in ϕ or R_g lead to an increase in the excess chemical potential, whereas an increase in R_g^{crd} causes a decrease in the excess chemical potential. Data reported in Fig. 2e are fitted globally to Eqs. S4 and S5 with a different $\left\langle R_{g0}^2\right\rangle^{1/2}$ for each IDP and a single effective R_g^{crd} as an adjustable parameter shared by all four proteins (Fig. S6).

A modification of the classic SPT suggested by Minton (37) allows the polymeric nature of the IDP to be taken into account within the SPT framework. Here the unfolded state is represented as a Gaussian cloud where the average number density of residues can be described as a function of the distance from the center of the mass of the protein, r_p, as

$$\rho\left(r_p\right) = n\frac{3}{2\pi\left\langle R_g^2\right\rangle^{1/2}} e^{-\frac{3r_p^2}{2\left\langle R_g^2\right\rangle}} \qquad \text{(Eq. S6)}$$

For a system consisting of the protein and a single sphere in solution, r_p can be expressed in terms of the distance between the centers of mass of the two objects, r_{sep}, and the radius of the sphere, which in this case is the radius of the crowding agent, R_g^{crd} (see Fig. S3). The probability, P_0, that no chain segments of the IDP lie within the volume of an arbitrarily placed hard sphere is calculated as a function of R_g of the IDP,

$$\log P_0 = -\left(\frac{6}{\pi}\right)^{1/2} \frac{n}{R_g r_{sep}} e^{-\frac{3r_{sep}^2}{2R_g^2}} \int_0^{R_g^{crd}} e^{-\frac{3r^2}{2R_g^2}} \sinh\frac{3r_{sep}r}{R_g^2} r\,dr. \qquad \text{(Eq. S7)}$$

Rescaling all distances relative to the size of the crowding agent, R_g^{crd}, yields

$$r_{sep} = f_{sep}R_g^{crd} \qquad \text{(Eq. S8)}$$

$$R_g = f_g R_g^{crd} \qquad \text{(Eq. S9)}$$

$$r = xR_g^{crd}. \qquad \text{(Eq. S10)}$$

P_0 can be written in terms of the rescaled dimensions as

$$\log P_0 = -\left(\frac{6}{\pi}\right)^{1/2} \frac{n}{f_g f_{sep}} e^{-\frac{3f_{sep}^2}{2f_g^2}} \int_0^1 e^{-\frac{3x^2}{2f_g^2}} \sinh\frac{3f_{sep}x}{f_g^2} x\,dx. \qquad \text{(Eq. S11)}$$

By equating the co-volume of the Gaussian cloud and the crowding agent with the co-volume of two hard spheres, an equivalent effective hard-sphere radius, R_g^{eff}, is obtained for each R_g sampled by the IDP, leading to

$$\overline{R}^{eff} = \frac{R_g^{eff}}{R_g^{crd}} = 3\int_0^{\infty}\left(\left(1-P_0\left(f_{sep}\right)\right)f_{sep}^2 df_{sep}\right)^{1/3} - 1 \qquad \text{(Eq. S12)}$$

This rescaled $\overline{R}^{eff}$ can then be inserted in Eq. S5 (Fig. S7).

Further extensions of SPT. In efforts to go beyond the simple description of a fluid of hard spheres, different extensions of the SPT have been implemented. SPT equations have been revised to account for ellipsoidal, cylindrical and infinite rod-like particles (37, 66-68). None of these corrections for different shapes can account for the trends observed in our experimental results. More recently, Qin and Zhou (69) have approached the problem of crowding on IDPs by calculating the co-volume on the basis of explicit simulations of the disordered protein and of the crowding agent in isolation. Even though this post-processing approach captures the effects induced by compact crowders on disordered proteins in their simulations, an extension of this method to polymeric expanded crowders as those used in our experiments has not yet been implemented. Complications in applying this approach come from the difficulty of calculating the correct co-volume between two disordered systems as well as taking into account the change in volume of the polymeric crowder at high concentrations (see the discussion about semidilute regime in the section *Flory-Huggins theories*).

To account for interactions between the crowders and the protein (47), attractive free energy terms have been included in SPT (46, 48), which resulted in the successful description of simulated data (48). A similar approach can be implemented here by adding an attractive interaction between IDP and crowder in Eq. S4. However, even with the functional form suggested by Kim & Mittal (48), the energy parameter would need to be different for each polymer length and possibly for different concentrations to obtain a quantitative fit of our data.

Flory-Huggins theories. A single polymer chain in good solvent adopts swollen conformations and follows a scaling exponent of 3/5, i.e. $R_g^{crd} \sim P^{3/5}$, where P is the number of Kuhn segments of the polymer. Three different concentration regimes need to be distinguished for a polymer in solution: the dilute regime, where the polymer chains are not overlapping; a semidilute regime, where the chains start to overlap and entangle; and a dense regime, where the chains are highly packed (Fig. S1 inset). ϕ^* is the overlap concentration, which separates the semidilute from the dilute regime. ϕ^* can be defined as the concentration of polymer where the volume fraction of the polymer chains in solution is equal to the volume fraction of a single polymer chain, i.e.

$$\phi^* = \frac{Pb^3}{P^{9/5}b^3} = P^{-4/5}. \qquad \text{(Eq. S13)}$$

Flory argued that in concentrated solutions and melts, the polymers exhibit the length scaling of an ideal chain ($R \sim P^{1/2}$) (70). Let us consider the case of one long chain with N segments (the IDP) in a polymer melt of shorter chains with P segments (the crowder). For simplicity, the segment length, b, is assumed to be equal for N-chain and P-chains. By equating the chemical potentials of crowders inside and outside the volume pervaded by the long chain, a relation between the end-to-end distance, *R*, of the long N-chain and the number of segments of the P-chains can be obtained (39):

$$\frac{R}{Nb^2} - \frac{N^2b^3}{PR^4} - \frac{1}{R} = 0 \qquad \text{(Eq. S14)}$$

If $P << N$, the third term in the equation can be neglected, and the size of the long chain can be described by the equation

$$R = N^{3/5}P^{-1/5}b. \qquad \text{(Eq. S15)}$$

In the case of large *P*, the second term in Eq. S14 can be neglected, and ideal scaling is recovered:

$$R = N^{1/2}b. \qquad \text{(Eq. S16)}$$

The crossover between small and large *P* is determined by equating Eqs. S15 and S16, which results in the threshold given by the Flory criterion: $P=N^{1/2}$ (39).

In our experiments, we investigate a **ternary system** composed of

- individual test chains with *N* segments (the IDP),
- a volume fraction, ϕ, of polymer chains with *P* segments (the crowding agent),
- and the solvent.

Similar to the case of the polymer melt (39), it is possible to describe the interaction between the long chain (the IDP) and the other polymers (the crowding agent) in terms of an effective medium interaction parameter. This term is obtained in the mean field approach of Joanny et al. (39) by equating the chemical potentials of the short chains (crowders) inside and outside the long chain (IDP). The effective medium interaction parameter is then given by:

$$u = \frac{1}{1+P\phi} \qquad \text{(Eq. S17)}$$

To study the effects of the crowder concentration on the size of the IDP, it is helpful to treat the IDP as a sequence of blobs of size R_P. In dilute solution, on length scales smaller than R_P, the *N*-chain behaves in the same way as the *P*-chains, whereas on length scales greater than R_P, the ternary properties of the system become relevant. In the latter case, the long chain formed by *N*/*P* blobs experiences the effective medium interaction that in rescaled units is (39)

$$u = \frac{1}{1+\phi/\phi^*}. \qquad \text{(Eq. S18)}$$

For a chain in good solvent, the radius of gyration is

$$R_g = bN^{3/5}u^{1/5} \qquad \text{(Eq. S19)}$$

And thus, in blob-rescaled units,

$$R_g = R_g^{crd}\left(\frac{N}{P}\right)^{3/5}\left(\frac{1}{1+\phi/\phi^*}\right)^{1/5} = R_{g0}\left(\frac{1}{1+\phi/\phi^*}\right)^{1/5} \qquad \text{(Eq. S20)}$$

The equation can be further extended to the semidilute regime (45). However, since this theory is only valid for the case of a long chain in a solution of shorter chains, this regime will only be approached marginally in our experiments.

For the data analysis presented in Figs. 3 and 4, we modified Eq. S20 by introducing a fitting parameter a, similar to the effective interaction term proposed by Nose (45):

$$R_g = R_{g0}\left(\frac{1}{1+a\phi/\phi^*}\right)^{1/5} \qquad \text{(Eq. 1a, see main text)}$$

All other parameters are not adjustable: ϕ^* is given by Eq. S13, and R_{g0} is obtained from the radius of gyration of each protein in the absence of crowding. a corresponds to a correction of the effective medium interaction term in Eq. S17, which is calculated for a very long chain in a bath of shorter chains, where the long chain is in good solvent and assumed to be large enough to accommodate the small chains (Fig. S8). The degree of expansion of the IDPs investigated here depends on their specific sequence (27), resulting in differences in the interaction term. The values of a obtained from the fits reflect the expected trend and exhibit an increasing deviation from $a = 1$ with increasing compactness of the protein (Table S4).

Extending the classical Flory-Huggins theory to the case of $P > N^{1/2}$ requires a realistic estimate of the density fluctuations in the solution when the solution is no longer dilute. A corresponding quantitative description is provided by the renormalized Flory-Huggins theory derived by Schäfer and Kappeler (44). Renormalization group theory establishes equivalence between microscopically different systems through scaling laws. In doing so, it allows to map a system of long chains (where common perturbation theories break down) onto a system of effectively short chains (where perturbation theories hold).

Here we introduce a length l_0, which defines a sort of lattice unit for the system and will be used to describe both the *N*-chain (the IDP) and the *P*-chains (the crowding agent). Following the treatment of Schäfer, we set l_0 equal to the length of the Kuhn segment of the *N*-chain. Here we adopted the same l_0 for all four proteins, according to the value obtained for ProTα. All the other chemical differences in structure and flexibility between the two chains will be absorbed in specific parameters. The renormalization is then introduced via a renormalized length, $l_R = l_0 / \lambda$, where $\lambda < 1$. Similarly, the number of segments of each chain and the volume fraction are rescaled as a function of the same λ parameter:

$$N \to N_R(\lambda) \qquad \text{(Eq. S21a)}$$

$$P \to P_R(\lambda) \qquad \text{(Eq. S21b)}$$

$$\phi \to \phi_R(\lambda)\,. \qquad \text{(Eq. S21c)}$$

The connection between the microscopic information and the renormalized parameter is given by:

$$1 = \frac{1}{N_R} + 2\left(\hat{u} B_N \phi l_0^{-3} N^{1.76} + \hat{u} \left(\frac{B_P}{B_N} \right)^{\frac{1}{0.588}} B_N^3 \phi l_0^{-3} N^{0.76} \right) N^{-0.76}, \qquad \text{(Eq. S22)}$$

where $\hat{u} = 5.756$ is a parameter connected to the overlap probability of chains in the semidilute regime; $B_i = \frac{R_{g0}^{(i)}}{0.8^{0.5} i^{0.588}}$, with $i = \{P, N\}$, contains the structural details of the single *N*- and *P*-chains in highly dilute solutions under the assumption that both *N*- and *P*-chains are in good solvent; and λ is chosen in order that $l_R = B_N \left(\frac{N}{N_R} \right)^{0.588}$. The radius of gyration is then given by

$$R_g(N, P, \phi, s_{NP}) = l_R^2 N_R \left\{ 0.636 + 0.165 N_R^{1/2} - 0.292 N_R^{1/2} f_{NP}^2 G(\phi_R P_R, \frac{N_R}{P_R}) \right\} \qquad \text{(Eq. S23)}$$

where $G(W, y) = W \int_0^\infty z^{-1/2} \frac{B(z) D(zy)}{1 + W D(zy)} dz$ is formally similar to the result obtained with the uniform expansion model (71) when the expansions due to the renormalization are condensed in terms of

$$B(z) = \frac{1}{6} e^{-z} - 2D(z) - \frac{8}{z}(D(z) - 1) - \frac{10}{z^2}(D(z) - 1 + \frac{z}{3}) \qquad \text{(Eq. S24)}$$

$$D(z) = \frac{2}{z^2}(e^{-z} - 1 + z). \qquad \text{(Eq. S25)}$$

The term f_{NP} is obtained under the assumption that interactions between *N*- and *P*-chains are small (44), leading to the equation $f_{NP}(s_{NP}) = 1 + \left(\frac{l_R}{s_{NP}} \right)^{-0.40}$, where s_{NP} is an invariant parameter with the dimensions of a length that quantifies the interactions between the *N*- and *P*-chains. f_{NP} can be connected to the Flory interaction parameter, χ, through the equation $|\chi| = |1 - f_{NP}|$. However, the parameter s_{NP} is preferred to χ since the latter is not invariant in the renormalization flow.

Note that in the global fit of the radii of gyration as a function of volume fraction of PEG for all IDPs (Fig. 3) with the renormalized Flory-Huggins theory (Fig. S9), the only adjustable parameter is s_{NP}, which is taken to be identical for all PEG sizes, but allowed to vary from protein to protein (Table S4); B_N and B_P are well-defined by experimental observables: B_N is calculated from the radius of gyration measured for each protein in the absence of crowding and B_P from the reported values of the radii of gyration for PEG (see Fig. S1 and *PEG scaling law*). If s_{NP} is close to zero, protein and crowding agent are indistinguishable in terms of inter- and intramolecular interactions. In Fig. S4a, the robustness of the functional form of the fitting function at different values of the fitting parameter s_{NP} is illustrated.

s_{NP} provides a new opportunity to quantify the interactions between protein and crowder beyond simple excluded volume effects, sometimes referred to as “chemical interactions”. The two variants of ProTα are well fitted with almost identical values of s_{NP} (see Table S4), whereas slightly different values were obtained for ACTR and IN, indicating small variations in the interactions between PEG and proteins. The values obtained are of the same order of magnitude as those previously reported in the literature for ternary solutions of synthetic polymers (72). A conversion of s_{NP} to the more intuitive parameter χ yields repulsive interactions in the range of 0.1-0.3 $k_{\mathrm{B}}T$ per segment between PEG and ProTα or ACTR. Interestingly, the fitted values for the experiments with ProTα and other polymers (see Table S5) suggest lower or even undetectable chemical interactions between the protein and the polymers. The contribution to the overall effects observed here is small (Fig. S4), but this trend is consistent with previous suggestions that PEG may be less inert than other polymers commonly employed in crowding experiments (13). One of the strengths of the approach presented here is the possibility to model a complex polymeric solution including such repulsive (or attractive) interactions.

In the application of the Flory-Huggins theories, we have approximated the length of a Kuhn segment by 0.76 nm for both protein and PEG according to previously reported persistence lengths (27, 73), which is equivalent to two bond segments. Consequently, the number of Kuhn segments of the *N*- and *P*-chains is given by half the degree of polymerization. In the case of PEG, the concentration in weight fraction was converted to volume fraction according to $\phi_{v/v} = \phi_{w/w} / (\rho(1-\phi_{w/w}) + \phi_{w/w})$ where $\rho = 1.12$ g/cm^3 is an average density for pure solutions of short PEGs. This approach is justified since no significant volume contraction is reported for solutions with the PEG concentrations used here. For the other crowding agents, where specific densities in solution were not reported, a direct correspondence between volume and weight fraction based on the densities of the pure substances was assumed.

PEG scaling law. According to Devanand *et al.* (29), water is a good solvent for PEG. Fig. S1 shows the scaling law obtained in (29) for long PEGs with the radii of gyration of PEGs used in the current study. The radii of gyration of PEGs are taken from (74, 75). In cases where only the hydrodynamic radius was determined experimentally, a conversion between hydrodynamic radius and radius of gyration was applied according to the ratio determined experimentally in (29). A deviation from the predicted scaling behavior (29) is visible for short PEGs, for which finite length effects start to dominate.

Estimation of the radius of gyration of IDPs at the Θ-state. An estimate of the radius of gyration for the four proteins in Θ-state conditions is obtained according to the previous analysis of the scaling exponent of disordered and unfolded sequences presented by Hofmann *et al.* (27), where the radius of gyration is linked to the scaling exponent by (76)

$$R_{g\Theta} = \sqrt{\frac{2 l_p^* b}{(2\nu+1)(2\nu+2)}} N^{\nu}, \qquad \text{(Eq. S26)}$$

where l_p^* = 0.4 nm, $b = 0.38$nm, ν□= 0.5, and N is the number of amino acids of the respective protein. The resulting radii of gyration in Θ-solvent are about 1.7 nm for IN, 1.8 nm for ProTα−C and ProTα−N, and 2 nm for ACTR.

Estimation of scaling exponents of IDPs. Following the empirical relation for the scaling exponent as a function of the hydrophobicity and net charge of unfolded and disordered proteins obtained previously (27), we estimated the scaling exponent for all full-length disordered sequences longer than 25 amino acids deposited in the Disprot database (v6.01) (77).The exponent ν is calculated according to

$$\nu(Q) = 1/3 + a\left[1 + \exp(x_0 - Q)/z\right]^{-1} \text{ and } \nu(H) = 1/3 + a\left[1 + \exp(x_0 + cH - d)/z\right]^{-1}, \quad \text{(Eq. S27)}$$

where H is the hydrophobicity according to the scale of Kyte and Doolittle (78), Q is the mean net charge of the sequence, $a = 0.394$, $z = 0.09$, $x_0 = 0.114$, $c = 1.72$ and $d = 0.9$.
The exponent is determined according to

$$\nu = \begin{cases} \nu(Q) & u^* > 0 \vee f = 0 \vee g = 0 \\ \nu(H) & u^* \leq 0 \vee f = 0 \wedge g = 0 \end{cases}, \quad \text{(Eq. S28)}$$

where f and g are the fractions of positive and negative charges in the sequence, respectively, and u^* is calculated according polyampholyte theory (11) as

$$u^* = \frac{4\pi l_b (f-g)^2}{\kappa^2} - \frac{\pi l_b^2 (f+g)^2}{\kappa}, \quad \text{(Eq. S29)}$$

where l_b is the Bjerrum length, and κ^{-1} is the Debye length.

We note that a higher percentage of globule-like IDPs has been estimated from simulations for a different subset of the Disprot database (33, 34). However, taking into account polyampholyte effects and patterning, it has been predicted that the majority of IDPs will maintain coil-like properties (34), and consequently they should be susceptible to the effect of crowding described here.

Physiological concentrations of IDPs in the nuclear pore complex and in RNA granules. A nuclear pore complex contains approximately 200 disordered nucleoporins containing FG-repeats, each with a length of ~600 amino acids (79, 80). From the chain length of these sequences, the overlap concentrations can be estimated to be in the range between of volume fractions between 0.05% and 4%. Assuming the nuclear pore to have a diameter of 30 nm and a height of 40 to 80 nm (25, 80, 81), the volume fraction occupied by the disordered nucleoporins is easily estimated as $\phi_{NP} = \frac{v_{NP}}{V_{pore}} = \frac{n_{NP} \cdot N_{NP} \cdot \langle v_{\mathrm{aa}} \rangle}{\pi r_{pore}^2 h_{pore}}$, where v_{NP} is the volume occupied by the nucleoporins; V_{pore} is the volume available in the pore; n_{NP} is the number of nucleoporins; N_{NP} is the sequence length of nucleoporins; $\langle v_{aa} \rangle$ is the average volume of the amino acid residues (approximately 0.13 nm^3); r_{pore} and h_{pore} are, respectively, the radius and the height of the pore. The volume fraction occupied by disordered nucleoporins is therefore between 25 and 55%, about an order of magnitude higher than the overlap concentration. The effects discussed in the main text are thus highly likely to be of importance for the conformational distributions of nucleoporins *in vivo*.

Various IDPs have been identified in RNA granules (22, 82, 83). Recent work has shown that the multivalency of these proteins can control phase separation and therefore the assembly of RNA granules (23). Even though the mechanism is not fully understood, and the proteins can undergo different conformational changes during the phase-separation process, the mechanism suggested by Li et al. (23) can be taken as an example to test whether the polymeric nature of the proteins is expected to be important in that range of concentrations. For the case of the engineered proline-rich motives (PRM) (23), the sequence is likely to be almost completely disordered. The molar concentration of the protein at which a given volume is occupied entirely by those IDPs is given by $c = 1/(N\langle v_{aa} \rangle \cdot N_A)$, where N is the number of amino acids of the sequence, and N_A is the Avogadro number. A calculation for the case of a sequence with 250 amino acids as those considered in the work of Li *et al. (23)* result in a local protein concentration in the range of ~50 mM. The corresponding overlap concentration is estimated to be between volume fractions of 1% and 6% or 0.5 and 3 mM. Considering that two different proteins are mixed in similar ratios in these experiments, a concentration of 0.25 to 1.5 mM is sufficient to reach the overlap regime. Phase separation for proteins of this length occurs at concentrations of approximately 50 μM, only 5 times less than the overlap concentration. However, due to phase separation, in the droplets, a concentration of proteins 100 times higher than the bulk solution is reported (23). The confinement in the droplet is therefore plausibly causing an increase of protein concentration significantly higher than the overlap concentration, and the overlap between disordered coils will affect the conformations of the disordered sequences. FUS and hnRNPA2, two disordered proteins identified in RNA granules, have been shown to exhibit a sol-gel transition *in vitro* (82) at a concentration above 1 mM with an overlap concentration ranging between 0.15 and 1 mM. Therefore, these proteins are expected to be in the semidilute regime before gelation.

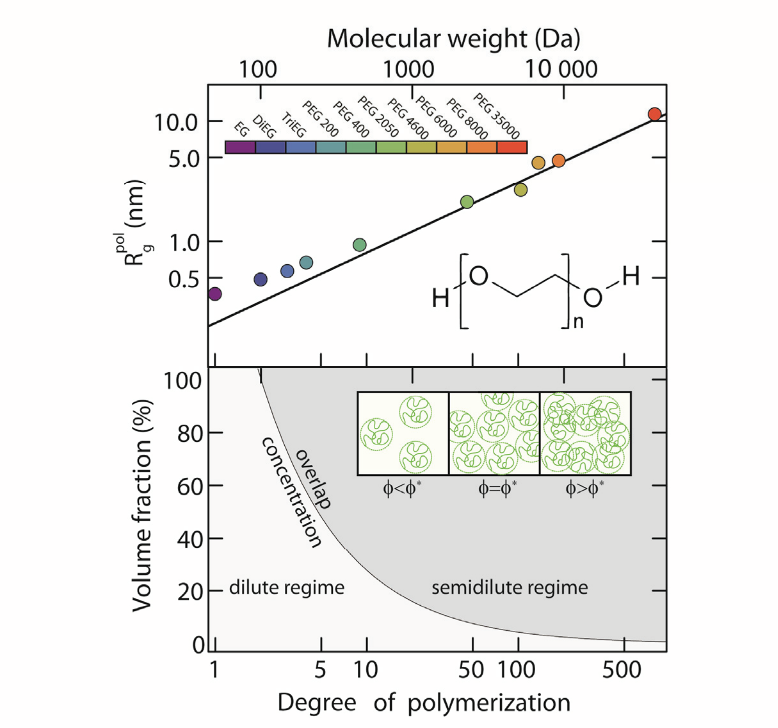


Figure S1. Polymer properties of polyethylene glycol. (a) Radius of gyration of PEG as a function of molecular weight/degree of polymerization (74, 75) with a fit (black line) to the scaling law $R_g^{pol} = 0.21\,\text{nm} \cdot P^{0.583}$ (26). The scaling exponent indicates that water is a good solvent for PEG. Deviations from the fit (which was obtained for PEG molecules over the entire range of lengths originally reported (29)) are due to finite length effects for small values of *P*. (b) Overlap concentration as obtained from Eq. S10, and schematic representation of a polymer solution in the dilute regime, at the overlap concentration, and in the semidilute regime.

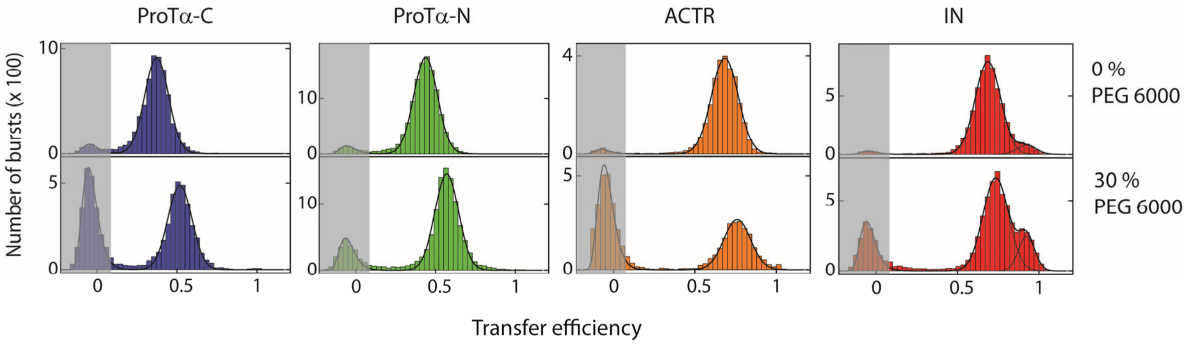


Figure S2. FRET efficiency histograms obtained with FRET dyes shifted toward longer wavelengths compared to Fig. 2. Histograms for ProTα-C, ProTα-N, ACTR, and IN variants labeled with ATTO546 and ATTO647N in the absence and presence of high PEG concentration. The shift towards higher excitation and emission wavelengths reduces the contribution of fluorescence background from impurities in the PEG and provides additional evidence that the peak broadening at high PEG concentration observed in Fig. 2 is mainly due to impurities in the solution. Only in the case of IN, a second peak corresponding to the folded state is detected, consistent with the observations and data analysis in Fig. 2. Gaussian and lognormal distributions were used to fit the peaks (solid lines). The donor-only peaks originating from molecules lacking an active acceptor dye are shaded in grey.

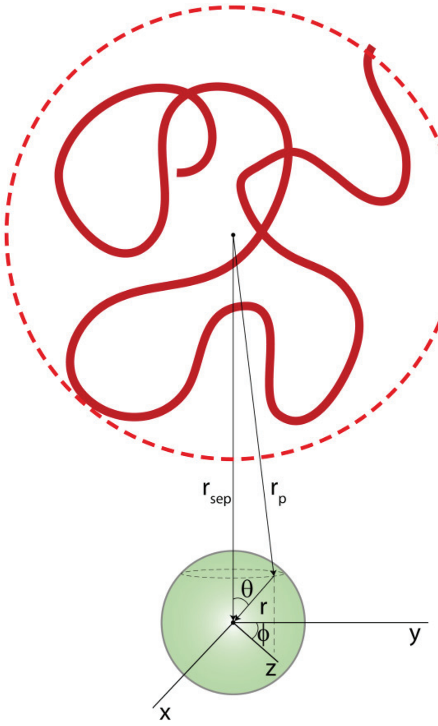


Figure S3. Gaussian cloud model. Representation of the Gaussian Cloud with a disordered conformation of the protein (in red), with the center of mass positioned at distance r_{sep} from a hard sphere (in green) with radius *r*. Adapted from (37).

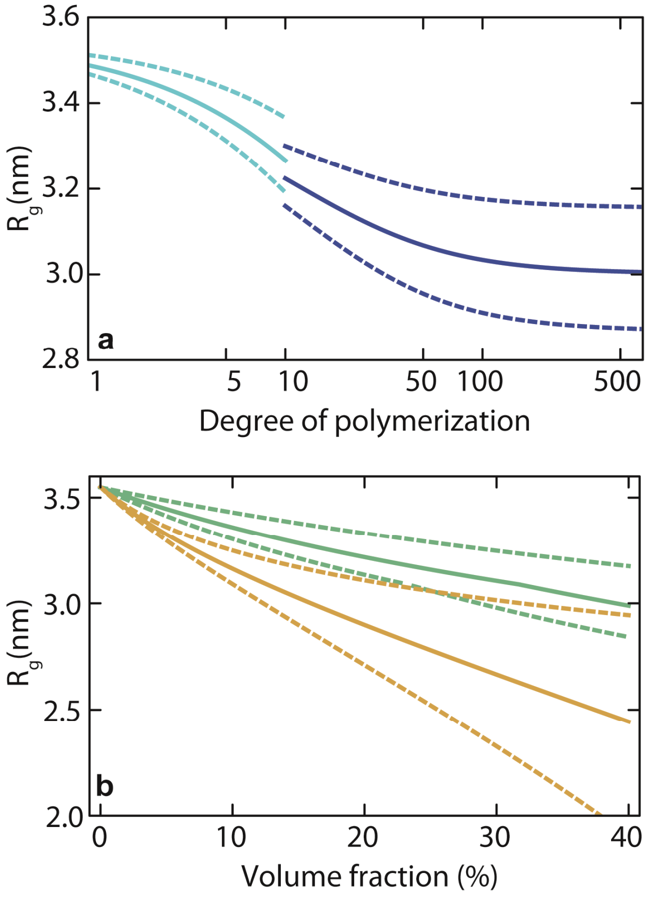


Figure S4. Effects of model parameters in the Flory-Huggins theories. Upper panel: Calculated radius of gyration of ProTα-C as a function of the degree of polymerization of PEG at a volume fraction of 15% according to Flory-Huggins theory (cyan) and renormalized Flory-Huggins theory (blue). Dashed curves show the change in the prediction of Flory-Huggins theory if a deviation of ±50% from the fitted value for the parameter *a* is assumed, and the response of renormalized Flory-Huggins theory to the fitting parameter s_{NP} between 0 (no interactions) and the upper limit of 0.1 nm (strong interactions). Lower Panel: estimation of the radius of gyration of ProTα-C at different volume fractions for PEG 400 (green solid curve) and PEG 6000 (yellow curve). Dashed lines report the response to the fitting parameter of the Flory-Huggins theory for PEG 400 and of the renormalized Flory-Huggins theory for PEG 6000 at the same conditions described above.

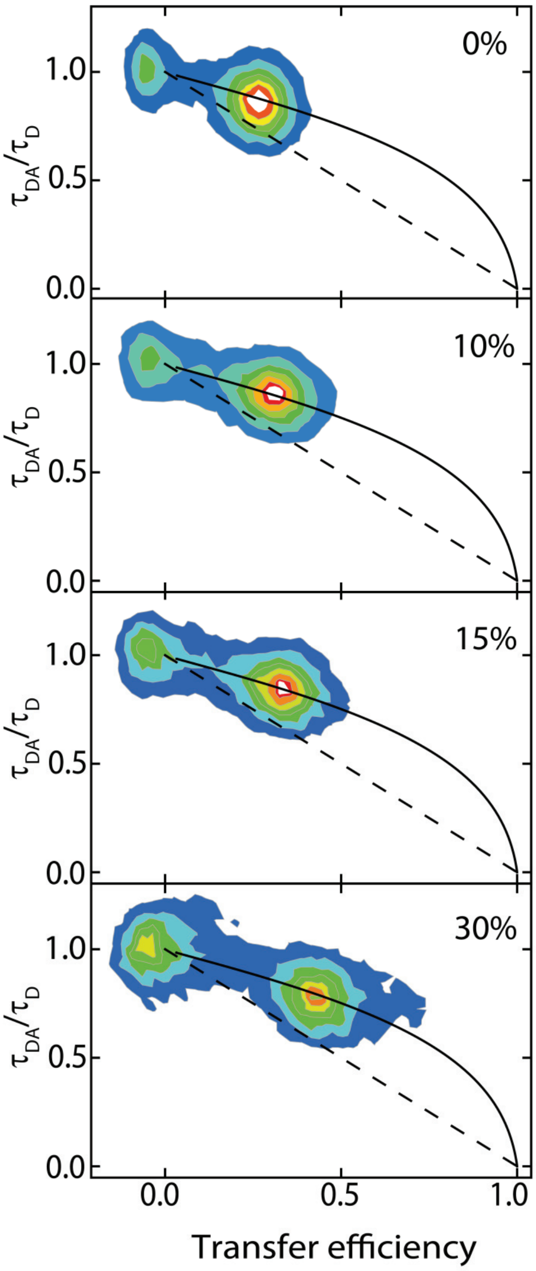


Figure S5. Multiparameter single-molecule fluorescence analysis. Two-dimensional histograms of relative donor τ_{DA}/τ_D lifetime versus FRET efficiency measured between 0 and 30% volume fraction of PEG 6000 compared to the expected trend for a fixed distance (black dashed line) and for a chain reconfiguring over the distribution of distances *P(r)* given by Eq. S2 (black solid line), as described in detail in the section *Fluorescence lifetimes and anisotropies*.

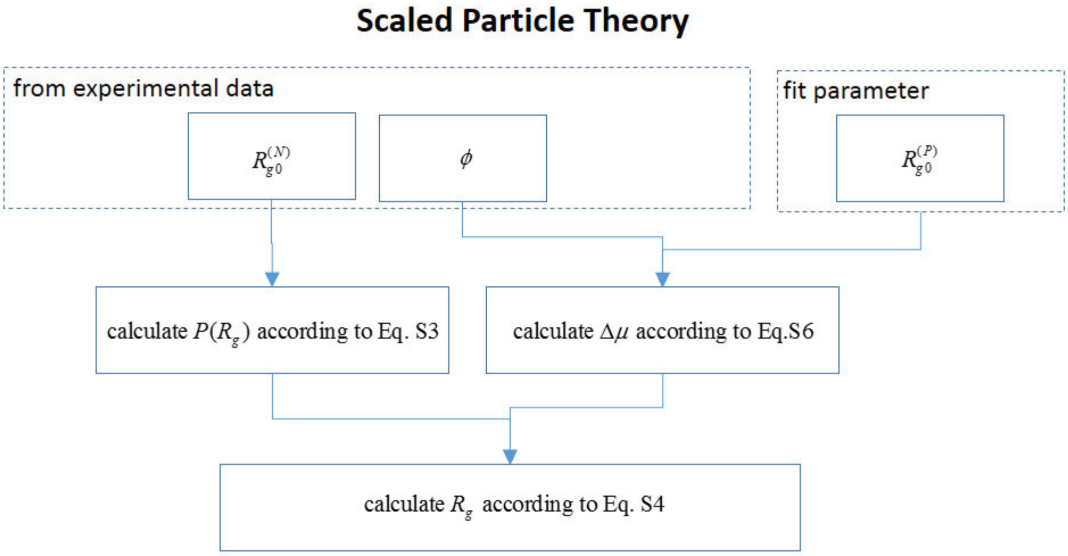


Figure S6. Flowchart for the fitting procedure with scaled-particle theory

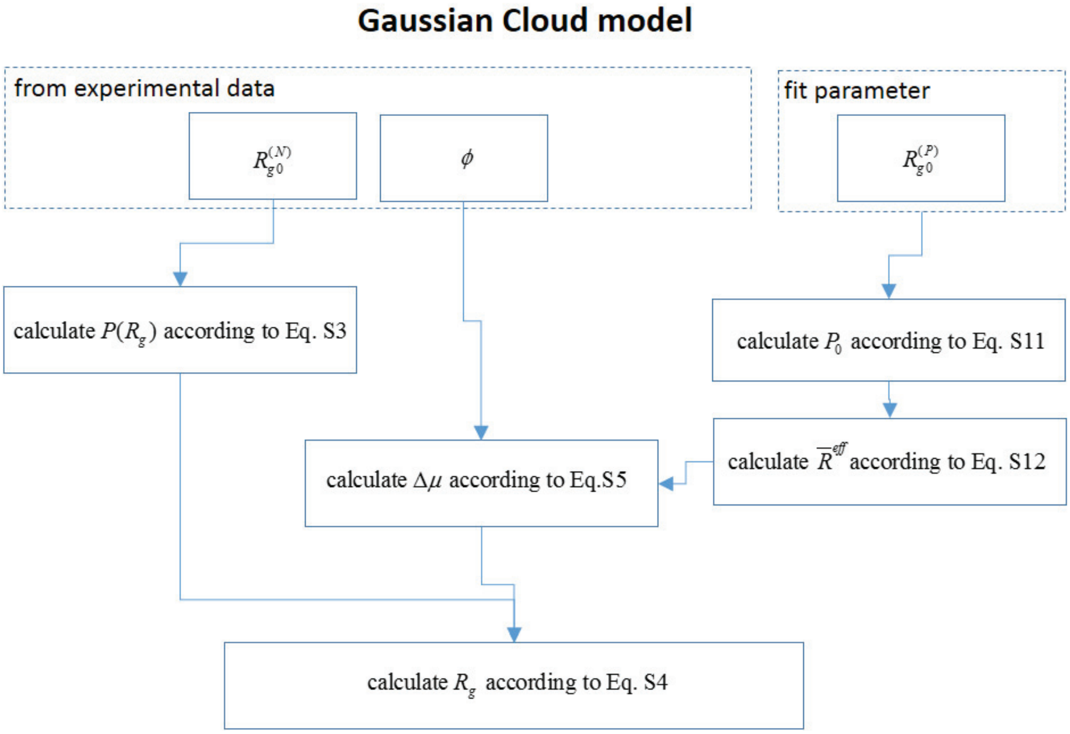


Figure S7. Flowchart for the fitting procedure with the Gaussian cloud model

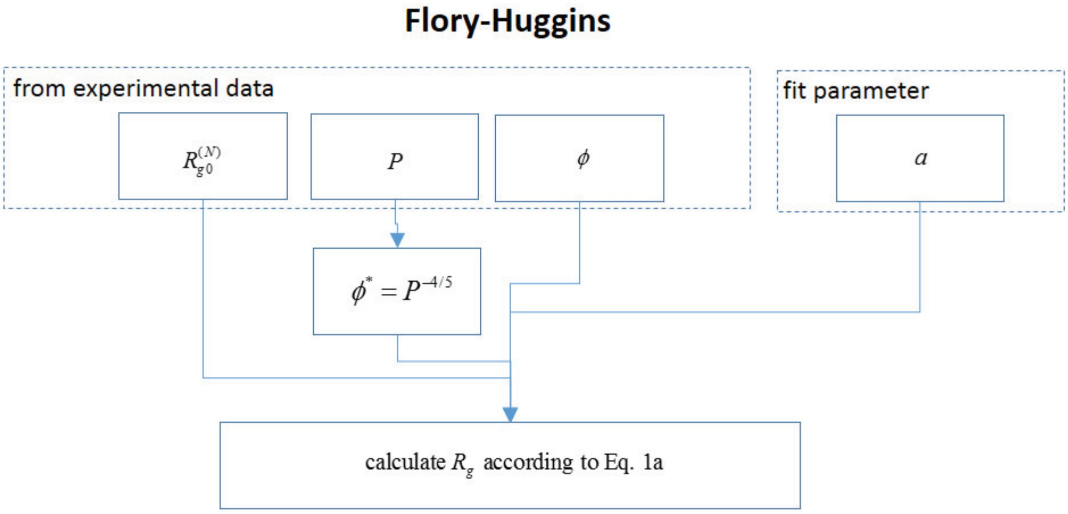


Figure S8. Flowchart for the fitting procedure with Flory-Huggins theory.

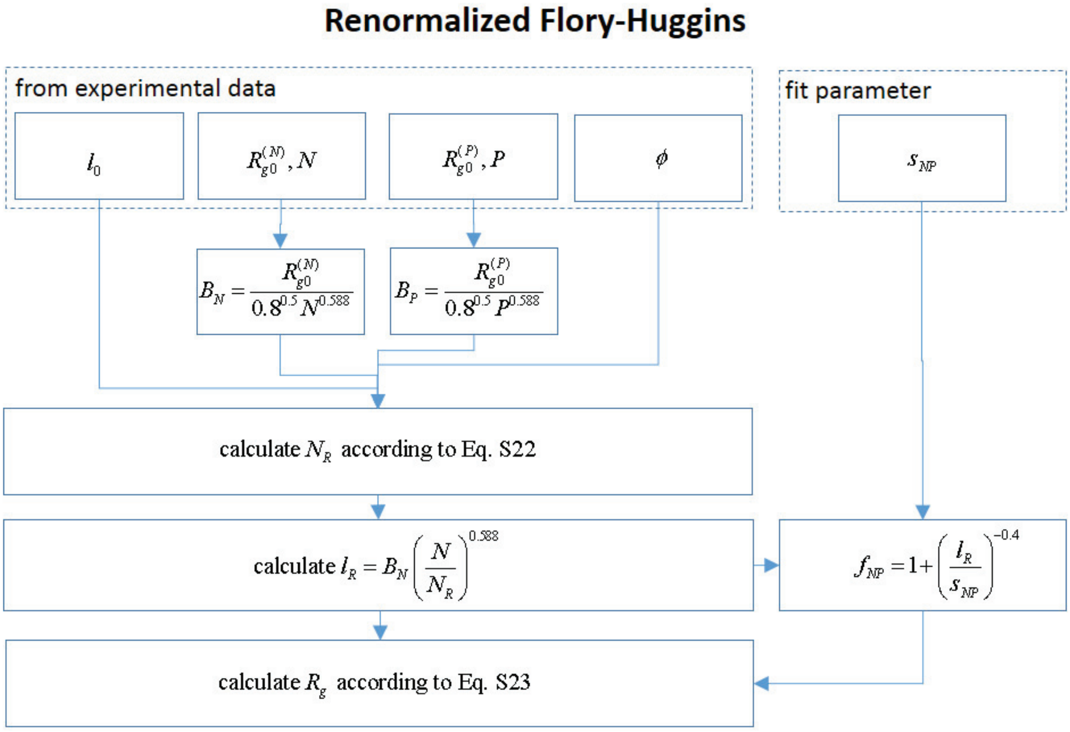


Figure S9. Flowchart for the fitting procedure with the renormalized Flory-Huggins theory.

	Mol. Weight (Da)	Average degree of polymerization
Ethylene glycol*	62.07	1
Diethylene glycol*	106.12	2
Triethylene glycol*	150.17	3
Polyethylene glycol 200*	190 - 210	4
Polyethylene glycol 400*	380 - 420	9
Polyethylene glycol 2050*	1900 - 2200	46
Polyethylene glycol 4600*	4400 - 4800	104
Polyethylene glycol 6000*	5000 - 7000	136
Polyethylene glycol 8000**Error! Bookmark not defined.**	7300 - 9000	185
Polyethylene glycol 35'000*	35000	795
Polyvinyl alcohol 10'000*	9000 - 10000	216
Polyvinyl alcohol 40'000*	31000 - 50000	920
Polyvinyl alcohol 90'000*	89000 - 98000	2125
Polyvinylpyrrolidone K90*	360000	3243
Polyvinylpyrrolidone 1.3M*	1300000	11712
Dextran 6000*	6000	37
Dextran 40'000*	40000	247
Dextran 100'000*	100000	617

*Sigma-Aldrich (Switzerland), **Error! Bookmark not defined.**Carl Roth (Germany)

Table S1. Crowding agents used in this study.

ProTα-C (C56-C110)	1 10 20 30 40 50 56 60 70 80 90 100 GP SDAAVDTSSE ITTKDLKEKK EVVEEAENGR DAPANGNAEN EENGEQEADN EVDEE**C**EEGG EEEEEEEEGD GEEEDGDEDE EAESATGKRA AEDDEDDDVD 110 TKKQKTDED**C**
ProTα-N (C2-C56)	1 10 20 30 40 50 56 60 70 80 90 100 GP **C**DAAVDTSSE ITTKDLKEKK EVVEEAENGR DAPANGNAEN EENGEQEADN EVDEE**C**EEGG EEEEEEEEGD GEEEDGDEDE EAESATGKRA AEDDEDDDVD 110 TKKQKTDEDD
ACTR (C1-C73)	1 10 20 30 40 50 60 70 73 GP **C**GTQNRPLLR NSLDDLVGPP SNLEGQSDER ALLDQLHTLL SNTDATGLEE IDRALGIPEL VNQGQALEPK QD**C**
IN (C8-C57)	1 10 20 30 40 50 57 GSH MFLDGID**C**AQ EEHEKAHSNF RAMASDFNLP PVVAKEIVAS CDKCQLKGEA MHGQVD**C**

Table S2. Sequences of the proteins used in this study.

	ProTα-C	ProTα-N	ACTR	IN
R_{g0} (nm)	3.46±0.02	3.03±0.02	2.47±0.02	1.95±0.02

	PEG 6000
R_g^{crd} (nm)	5.8±0.1

Table S3. Global fit of PEG 6000 data with scaled-particle theory. Parameters obtained from fitting the data in Fig. 2 with Eq. S4

	$P < N^{1/2}$	$P > N^{1/2}$
	a	s_{NP} (nm)
ProTα-C	1.18±0.05	0.032±0.004
ProTα-N	0.7±0.1	0.024±0.008
ACTR	0.8±0.1	0.08±0.02
IN	0.05±0.02	$< 2 \cdot 10^{-6}$ (*)

(*) s_{NP} is sufficiently close to zero that the ternary system reduces to a binary system where the protein and the crowding agents cannot be distinguished

Table S4. Global fit of Protα-C, Protα-N, ACTR and IN in the presence of PEG with Flory-Huggins theories. Parameters obtained from fitting the complete dataset in Fig. 3 with Eqs. 1a and 1b.

	s_{NP} *(nm)*
PEG	0.032±0.004
PVA	$< 4\cdot10^{-5\,(*)}$
PVP	$< 4\cdot10^{-6\,(*)}$
Dextran	$< 3\cdot10^{-5\,(*)}$

(*) s_{NP} is sufficiently close to zero that the ternary system reduces to a binary system where the protein and the crowding agents cannot be distinguished

Table S5. Fit results with Flory-Huggins theories of ProTα-C collapse in presence of PVA, PVP and Dextran. Parameters obtained from fitting the datasets of ProTα-C in the presence of PEG (Fig. 3), PVA, PVP and Dextran (Fig. 5) with Eq. 1b.

Supporting Information references

1. Dyson HJ & Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* 6(3):197-208.
2. Dunker AK, Silman I, Uversky VN, & Sussman JL (2008) Function and structure of inherently disordered proteins. *Curr Opin Struct Biol* 18(6):756-764.
3. Jensen MR, Ruigrok RW, & Blackledge M (2013) Describing intrinsically disordered proteins at atomic resolution by NMR. *Curr. Opin. Struct. Biol.*
4. Wright PE & Dyson HJ (2009) Linking folding and binding. *Curr. Opin. Struct. Biol.*
5. Ferreon AC, Moran CR, Gambin Y, & Deniz AA (2010) Single-molecule fluorescence studies of intrinsically disordered proteins. *Methods in enzymology* 472:179-204.
6. Schuler B, Müller-Späth S, Soranno A, & Nettels D (2012) Application of confocal single-molecule FRET to intrinsically disordered proteins. *Methods Mol. Biol.* 896:21-45.
7. Ferreon AC, Ferreon JC, Wright PE, & Deniz AA (2013) Modulation of allostery by protein intrinsic disorder. *Nature* 498(7454):390-394.
8. Mao AH, Lyle N, & Pappu RV (2013) Describing sequence-ensemble relationships for intrinsically disordered proteins. *Biochem J* 449(2):307-318.
9. Lyle N, Das RK, & Pappu RV (2013) A quantitative measure for protein conformational heterogeneity. *J. Chem. Phys.* 139(12).
10. Fisher CK & Stultz CM (2011) Protein Structure along the Order-Disorder Continuum. *J. Am. Chem. Soc.* 133(26):10022-10025.
11. Müller-Späth S*, et al.* (2010) Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. U. S. A.* 107(33):14609-14614.
12. Uversky VN (2009) Intrinsically disordered proteins and their environment: effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein J.* 28(7-8):305-325.
13. Zhou HX, Rivas GN, & Minton AP (2008) Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.* 37:375-397.
14. Gershenson A & Gierasch LM (2011) Protein folding in the cell: challenges and progress. *Curr. Opin. Struct. Biol.* 21(1):32-41.
15. Dedmon MM, Patel CN, Young GB, & Pielak GJ (2002) FlgM gains structure in living cells. *Proceedings of the National Academy of Sciences of the United States of America* 99(20):12681-12684.
16. McNulty BC, Young GB, & Pielak GJ (2006) Macromolecular crowding in the Escherichia coli periplasm maintains alpha-synuclein disorder. *Journal of molecular biology* 355(5):893-897.
17. Munishkina LA, Cooper EM, Uversky VN, & Fink AL (2004) The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *J Mol Recognit* 17(5):456-464.
18. Szasz C*, et al.* (2011) Protein Disorder Prevails under Crowded Conditions. *Biochemistry* 50(26):5834-5844.
19. Hong JA & Gierasch LM (2010) Macromolecular Crowding Remodels the Energy Landscape of a Protein by Favoring a More Compact Unfolded State. *Journal of the American Chemical Society* 132(30):10445-10452.
20. Mikaelsson T, Aden J, Johansson LBA, & Wittung-Stafshede P (2013) Direct Observation of Protein Unfolded State Compaction in the Presence of Macromolecular Crowding. *Biophysical Journal* 104(3):694-704.

21. Johansen D, Jeffries CM, Hammouda B, Trewhella J, & Goldenberg DP (2011) Effects of macromolecular crowding on an intrinsically disordered protein characterized by small-angle neutron scattering with contrast matching. *Biophysical journal* 100(4):1120-1128.
22. Han TNW*, et al.* (2012) Cell-free Formation of RNA Granules: Bound RNAs Identify Features and Components of Cellular Assemblies. *Cell* 149(4):768-779.
23. Li PL*, et al.* (2012) Phase transitions in the assembly of multivalent signalling proteins. *Nature* 483(7389):336-U129.
24. Brangwynne CP (2011) Soft active aggregates: mechanics, dynamics and self-assembly of liquid-like intracellular protein bodies. *Soft Matter* 7(7):3052-3059.
25. Rout MP*, et al.* (2000) The yeast nuclear pore complex: Composition, architecture, and transport mechanism. *J. Cell. Biol.* 148(4):635-651.
26. Elcock AH (2010) Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. *Curr Opin Struc Biol* 20(2):196-206.
27. Hofmann H*, et al.* (2012) Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* 109(40):16155-16160.
28. Soranno A*, et al.* (2012) Quantifying internal friction in unfolded and intrinsically disordered proteins with single molecule spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* 109(44):17800-17806.
29. Devanand K & Selser JC (1991) Asymptotic-Behavior and Long-Range Interactions in Aqueous-Solutions of Poly(Ethylene Oxide). *Macromolecules* 24(22):5943-5947.
30. Zimmerman SB & Trach SO (1991) Estimation of Macromolecule Concentrations and Excluded Volume Effects for the Cytoplasm of Escherichia-Coli. *Journal of molecular biology* 222(3):599-620.
31. Harris JM (1992) *Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications* (Plenum Press, New York).
32. Uversky VN, Gillespie JR, & Fink AL (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* 41(3):415-427.
33. Mao AH, Crick SL, Vitalis A, Chicoine CL, & Pappu RV (2010) Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proceedings of the National Academy of Sciences of the United States of America* 107(18):8183–8188.
34. Das RK & Pappu RV (2013) Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. U S A* 110(33):13392-13397.
35. Ziv G & Haran G (2009) Protein Folding, Protein Collapse, and Tanford's Transfer Model: Lessons from Single-Molecule FRET. *J. Am. Chem. Soc.* 131(8):2942-2947.
36. Cheung MS, Klimov D, & Thirumalai D (2005) Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proceedings of the National Academy of Sciences of the United States of America* 102(13):4753-4758.
37. Minton AP (2005) Models for excluded volume interaction between an unfolded protein and rigid macromolecular cosolutes: Macromolecular crowding and protein stability revisited. *Biophys. J.* 88(2):971-985.
38. Mittal J & Best RB (2010) Dependence of protein folding stability and dynamics on the density and composition of macromolecular crowders. *Biophys. J.* 98(2):315-320.
39. Joanny JF, Grant P, Pincus P, & Turkevich LA (1981) Conformations of Polydisperse Polymer-Solutions - Bimodal Distribution. *J Appl Phys* 52(10):5943-5948.
40. Edwards SF (1966) Theory of Polymer Solutions at Intermediate Concentration. *P Phys Soc Lond* 88(560P):265-&.
41. Schäfer L (1999) *Excluded Volume Effects in Polymer Solutions as Explained by the Renormalization Group* (Springer, Berlin).

42. Tran HT & Pappu RV (2006) Toward an accurate theoretical framework for describing ensembles for proteins under strongly denaturing conditions. *Biophysical journal* 91(5):1868-1886.
43. Wilson KG (1983) The Renormalization-Group and Critical Phenomena. *Rev. Mod. Phys.* 55(3):583-600.
44. Schäfer L & Kappeler C (1993) Interaction Effects on the Size of a Polymer-Chain in Ternary Solutions - a Renormalization-Group Study. *Journal of Chemical Physics* 99(8):6135-6154.
45. Nose T (1986) Chain Dimension of a Guest Polymer in the Semidilute Solution of Compatible and Incompatible Polymers. *J Phys-Paris* 47(3):517-527.
46. Minton AP (2013) Quantitative assessment of the relative contributions of steric repulsion and chemical interactions to macromolecular crowding. *Biopolymers* 99(4):239-244.
47. Sarkar M., Li C., & Pielak GJ (2013) Soft interactions and crowding. *Biophys. Rev.* 5(2):187-194.
48. Kim YC & Mittal J (2013) Crowding Induced Entropy-Enthalpy Compensation in Protein Association Equilibria. *Phys. Rev. Lett.* 110(20).
49. Shoemaker BA, Portman JJ, & Wolynes PG (2000) Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc. Natl. Acad. Sci. U. S. A.* 97(16):8868-8873.
50. Trizac E, Levy Y, & Wolynes PG (2010) Capillarity theory for the fly-casting mechanism. *Proceedings of the National Academy of Sciences of the United States of America* 107(7):2746-2750.
51. Schreiber G, Haran G, & Zhou HX (2009) Fundamental Aspects of Protein-Protein Association Kinetics. *Chem Rev* 109(3):839-860.
52. Vareli K, Frangou-Lazaridis M, van der Kraan I, Tsolas O, & van Driel R (2000) Nuclear distribution of prothymosin alpha and parathymosin: Evidence that prothymosin alpha is associated with RNA synthesis processing and para-thymosin with early DNA replication. *Exp. Cell Res.* 257(1):152-161.
53. Enkemann SA, Ward RD, & Berger SL (2000) Mobility within the nucleus and neighboring cytosol is a key feature of prothymosin-alpha. *J. Histochem. Cytochem.* 48(10):1341-1355.
54. Gelman H, Platkov M, & Gruebele M (2012) Rapid perturbation of free-energy landscapes: from in vitro to in vivo. *Chemistry* 18(21):6420-6427.
55. Phillip Y, Kiss V, & Schreiber G (2012) Protein-binding dynamics imaged in a living cell. *Proc. Natl. Acad. Sci. U. S. A.* 109(5):1461-1466.
56. Sakon JJ & Weninger KR (2010) Detecting the conformation of individual proteins in live cells. *Nature methods* 7(3):203-205.
57. Gast K*, et al.* (1995) Prothymosin alpha: a biologically active protein with random coil conformation. *Biochemistry* 34(40):13211-13218.
58. Zheng R, Jenkins TM, & Craigie R (1996) Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. *Proceedings of the National Academy of Sciences of the United States of America* 93(24):13659-13664.
59. Demarest SJ*, et al.* (2002) Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* 415(6871):549-553.
60. Yi S, Brickenden A, & Choy WY (2008) A new protocol for high-yield purification of recombinant human prothymosin alpha expressed in Escherichia coli for NMR studies. *Protein expression and purification* 57(1):1-8.
61. Müller BK, Zaychikov E, Bräuchle C, & Lamb DC (2005) Pulsed interleaved excitation. *Biophys. J.* 89(5):3508-3522.

62. Sisamakis E, Valeri A, Kalinin S, Rothwell PJ, & Seidel CAM (2010) Accurate Single-Molecule FRET Studies Using Multiparameter Fluorescence Detection. *Methods Enzymol.* 475:455-514.
63. Strickler SJ & Berg RA (1962) Relationship between Absorption Intensity and Fluorescence Lifetime of Molecules. *Journal of Chemical Physics* 37(4):814-&.
64. Sherman E & Haran G (2006) Coil-globule transition in the denatured state of a small protein. *Proc. Natl. Acad. Sci. U. S. A.* 103(31):11539-11543.
65. Sanchez IC (1979) Phase-Transition Behavior of the Isolated Polymer-Chain. *Macromolecules* 12(5):980-988.
66. Minton AP (1998) Molecular crowding: analysis of effects of high concentrations of inert cosolutes on biochemical equilibria and rates in terms of volume exclusion. *Methods in enzymology* 295:127-149.
67. Boublik T (1974) Statistical Thermodynamics of Convex Molecule Fluids. *Mol. Phys.* 27(5):1415-1427.
68. Oversteegen SM & Roth R (2005) General methods for free-volume theory. *J. Chem. Phys.* 122(21).
69. Qin S & Zhou HX (2013) Effects of Macromolecular Crowding on the Conformational Ensembles of Disordered Proteins. *The journal of physical chemistry letters* 4(20).
70. De Gennes PG (1979) *Scaling concepts in polymer physics* (Cornell University Press, Ithaca, N.Y.).
71. Raphael E, Fredrickson GH, & Pincus P (1992) One Long-Chain among Shorter Chains - the Flory Approach Revisited. *Journal De Physique Ii* 2(10):1811-1823.
72. Kappeler C, Schafer L, & Fukuda T (1991) Light-Scattering from Ternary Polymer-Solutions - an Analysis of the Apparent Radius of Gyration. *Macromolecules* 24(10):2715-2718.
73. Ferry Kienberger*, et al.* (2000) Static and Dynamical Properties of Single Poly(Ethylene Glycol) Molecules Investigated by Force Spectroscopy. *Single Molecules* 1(2):123-128.
74. Kuga S (1981) Pore-Size Distribution Analysis of Gel Substances by Size Exclusion Chromatography. *J. Chromatogr.* 206(3):449-461.
75. Hosek M & Tang JX (2004) Polymer-induced bundling of F actin and the depletion force. *Phys. Rev. E* 69(5).
76. Hammouda B (1993) Sans from Homogeneous Polymer Mixtures - a Unified Overview. *Adv. Polym. Sci.* 106:87-133.
77. Sickmeier M*, et al.* (2007) DisProt: the database of disordered proteins. *Nucleic Acids Res.* 35:D786-D793.
78. Kyte J & Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157(1):105-132.
79. Milles S & Lemke EA (2011) Single molecule study of the intrinsically disordered FG-repeat nucleoporin 153. *Biophys. J.* 101(7):1710-1719.
80. Rout MP, Aitchison JD, Magnasco MO, & Chait BT (2003) Virtual gating and nuclear transport: the hole picture. *Trends Cell Biol.* 13(12):622-628.
81. Alber F*, et al.* (2007) The molecular architecture of the nuclear pore complex. *Nature* 450(7170):695-701.
82. Kato M*, et al.* (2012) Cell-free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels. *Cell* 149(4):753-767.
83. Kato Y & Nakamura A (2012) Roles of cytoplasmic RNP granules in intracellular RNA localization and translational control in the Drosophila oocyte. *Dev. Growth Differ.* 54(1):19-31.

3 Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells

Iwo König[1], Arash Zarrine-Afsar[1], Mikayel Aznauryan[1], Andrea Soranno, Bengt Wunderlich, Fabian Dingfelder, Jakob C Stüber, Andreas Plückthun, Daniel Nettels, and Benjamin Schuler

[1] These authors contributed equally to this work

(Nature Methods, 2015)

The author established the in-cell experiments, carried out the in-cell single-molecule FRET measurements and analyzed the data to obtain dimensions, translational diffusion times and chain reconfiguration times of the measured proteins.

Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells

Iwo König[1,3], Arash Zarrine-Afsar[1–3], Mikayel Aznauryan[1–3], Andrea Soranno[1], Bengt Wunderlich[1], Fabian Dingfelder[1], Jakob C Stüber[1], Andreas Plückthun[1], Daniel Nettels[1] & Benjamin Schuler[1]

Single-molecule methods have become widely used for quantifying the conformational heterogeneity and structural dynamics of biomolecules *in vitro*. Their application *in vivo*, however, has remained challenging owing to shortcomings in the design and reproducible delivery of labeled molecules, the range of applicable analysis methods, and suboptimal cell culture conditions. By addressing these limitations in an integrated approach, we demonstrate the feasibility of probing protein dynamics from milliseconds down to the nanosecond regime in live eukaryotic cells with confocal single-molecule Förster resonance energy transfer (FRET) spectroscopy. We illustrate the versatility of the approach by determining the dimensions and submicrosecond chain dynamics of an intrinsically disordered protein; by detecting even subtle changes in the temperature dependence of protein stability, including in-cell cold denaturation; and by quantifying the folding dynamics of a small protein. The methodology opens possibilities for assessing the effect of the cellular environment on biomolecular conformation, dynamics and function.

Advances in methodology are making it gradually more feasible to investigate biomolecular processes in their native cellular environment. The ultimate goal is to reach quantitative molecular understanding with the same rigor as in test-tube experiments. Owing to its sensitivity, fluorescence has become particularly popular—especially with the broad availability of fluorescent proteins[1]—for investigating cellular localization, biomolecular interactions[2], and protein stability and folding dynamics[3] in living cells. This sensitivity has enabled the use of single-molecule tracking[4,5], intracellular fluorescence correlation spectroscopy (FCS)[6], and image correlation methods[7] for investigating molecular diffusion and transport processes. Super-resolution methods now enable localization and dynamic imaging below the diffraction limit in live cells[8,9]. In prokaryotic cells, substantial progress has recently been reported for single-molecule detection *in vivo*, including the investigation of stochasticity in gene expression[10] and the mechanisms of DNA replication[11] and repair[12], for example, in combination with electroporation of labeled molecules[13]. The first single-molecule FRET experiments for detecting conformational changes of proteins in live eukaryotic cells have also been reported[14]. However, in spite of these advances, the use of single-molecule spectroscopy for resolving the subnanometer conformational changes of biomolecules and especially the wide range of relevant timescales has remained challenging *in vivo*.

This challenge largely originates from the multifactorial requirements for successful in-cell single-molecule measurements, such as (i) the cellular host system needs to be prepared in a fashion that minimizes autofluorescence; (ii) suitable excitation wavelengths and FRET dyes are needed that optimize molecular brightness of the sample compared to cellular background; (iii) the sample needs to be delivered into the cell in a targeted and reproducible fashion at the low concentrations suitable for single-molecule detection; (iv) measurements must follow sample delivery promptly enough to allow the dynamics of cellular targeting and localization to be followed and to avoid intracellular sample degradation; (v) intact cellular activity must be preserved and controlled for; (vi) both immobile and freely diffusing molecules should be detectable; and (vii) data analysis must make optimal use of the limited reservoir of molecules per cell to enable access to submillisecond timescales.

Here we show that with an integrated approach optimizing the points listed above, a surprisingly broad spectrum of confocal single-molecule FRET and nanosecond FCS methods that have previously been limited to *in vitro* experiments[15–18] could be used in live eukaryotic cells. As a result, even small conformational changes of protein molecules and conformational dynamics down to the nanosecond range became accessible in in-cell experiments. To illustrate the potential of the methodology, we investigated, in live cells, (i) the conformational distributions and nanosecond dynamics of an intrinsically disordered protein (IDP); (ii) the changes in the temperature-dependent conformational stability of a protein, including its cold denaturation; and (iii) the millisecond kinetics of a protein-folding reaction.

---

[1]Department of Biochemistry, University of Zurich, Zurich, Switzerland. [2]Present addresses: Techna Institute for the Advancement of Technology for Health, University Health Network, University of Toronto, Toronto, Ontario, Canada (A.Z.-A.); Keenan Research Centre for Biomedical Science, St. Michael's Hospital, Toronto, Ontario, Canada (A.Z.-A.); Interdisciplinary Nanoscience Center, Aarhus University, Aarhus, Denmark (M.A.). [3]These authors contributed equally to this work. Correspondence should be addressed to B.S. (schuler@bioc.uzh.ch).

RECEIVED 27 JANUARY; ACCEPTED 22 MAY; PUBLISHED ONLINE 6 JULY 2015; DOI:10.1038/NMETH.3475

RESULTS

In-cell confocal single-molecule FRET spectroscopy

Among the main considerations for performing intracellular single-molecule spectroscopy are, one, the reduction of cellular autofluorescence to a level that does not interfere with single-molecule detection and, two, the targeted application of labeled molecules into cells[5,10,19]. Autofluorescence in the visible wavelength range tends to decrease with increasing excitation wavelength[20], and above ~520 nm, autofluorescence of many cell lines is already sufficiently low. We chose to work with recombinantly produced proteins labeled via thiol-maleimide chemistry with organic fluorophores—in this case, Atto 532 in combination with three different acceptors: Biotium CF680R, Atto 647N and Abberior Star 635 (Online Methods).

We chose microinjection[2,5,14] for the targeted introduction of FRET-labeled proteins into cultured eukaryotic cells. Compared to transfection techniques such as scrape-loading[21], electroporation[22] and cell-penetrating peptides[23], microinjection offered the best control and reproducibility of the amount of protein applied; it enabled the shortest delay between protein application and measurement and is compatible with fluorescence detection directly after injection; it preserved cell adherence, which avoids problems with positional drift of the cells during measurements; and it creates the opportunity to inject molecules selectively into different cellular compartments. Microinjection enabled the reproducible introduction of sample concentrations in the picomolar to low nanomolar range, with cell viability extending far beyond the required observation times (Online Methods and **Supplementary Fig. 1**).

Measurements comprised the following iterative steps. First, we performed a reference measurement in the extracellular medium by placing the tip of the injection needle close to the confocal observation volume and ejecting several pulses of sample into free solution. Fluorescence bursts from molecules diffusing through the laser focus allowed us to confirm the integrity of the sample and functionality of the microinjection system. Analysis of these data permitted direct comparison of the conformational and dynamic properties of the labeled protein outside and inside a cell. In the next step, we injected a suitable amount of protein into a cell. The intracellular distribution of the labeled protein could then be monitored by a confocal sample scan (**Fig. 1** and **Supplementary Fig. 2**). Because we used single-photon counting in combination with pulsed excitation, fluorescence lifetime imaging microscopy[24] (FLIM) could be employed to differentiate

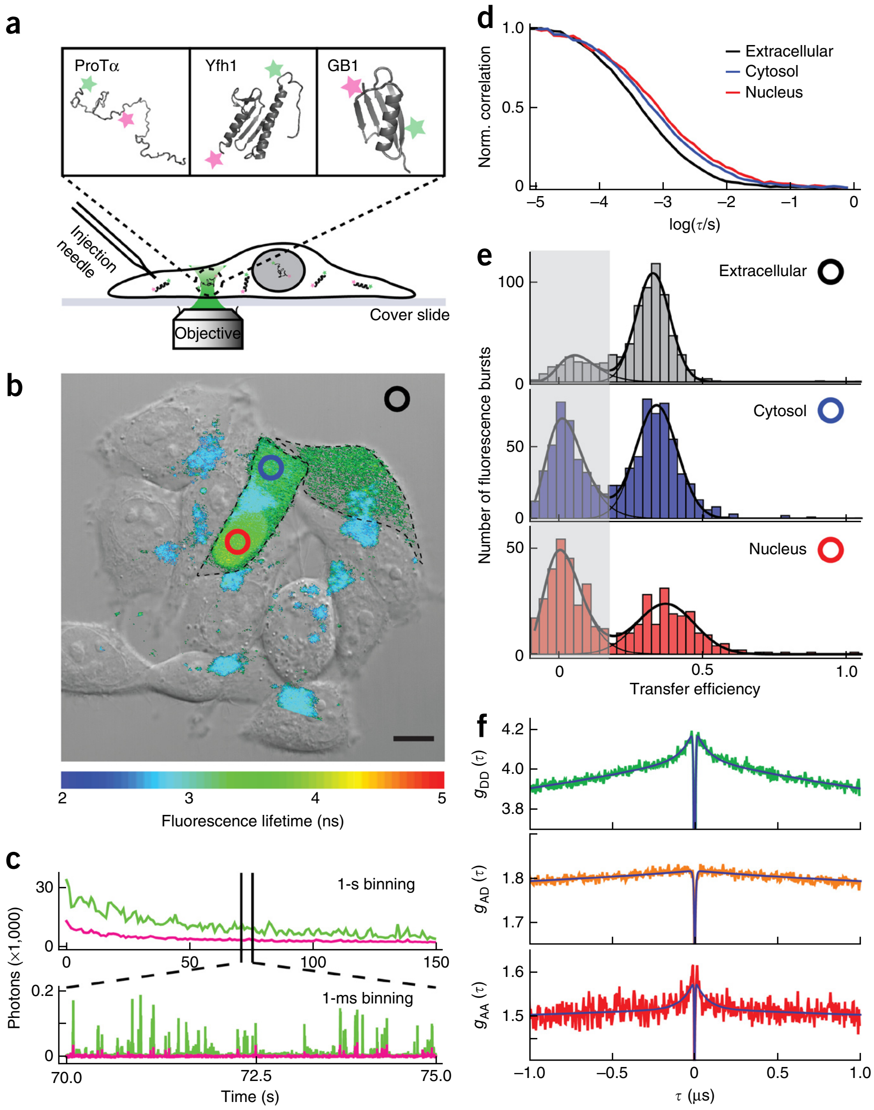


Figure 1 | Single-molecule fluorescence measurements in living cells. (**a**) Schematic of confocal single-molecule FRET studies of fluorescently labeled prothymosin α (ProTα), yeast frataxin homolog Yfh1, and IgG-binding domain of protein G (GB1) microinjected into adherent HeLa cells. (**b**) Fluorescence lifetime image (overlaid with differential interference contrast image) of HeLa cells after cytosolic injection of ProTα (injected cells indicated by dashed lines); donor fluorescence of ProTα exhibits a higher average lifetime (~3.5 ns) than the cellular autofluorescence (<3 ns). Circles indicate positions of the confocal volume in the nucleus (red), cytosol (blue) and extracellular medium (black). Scale bar, 10 μm. (**c**) Fluorescence time trace recorded in the cytosol of an injected cell. Photobleaching results in a characteristic intensity decay (1-s binning). With 1-ms binning, fluorescence bursts of donor (green) and acceptor (magenta) photons are visible. (**d**) Donor-acceptor cross-correlation curves were used to compare the translational diffusion of ProTα in buffer ($\tau_{\mathrm{Diff}} = 0.7 \pm 0.1$ ms, $n = 5$), cytosol ($\tau_{\mathrm{Diff}} = 1.8 \pm 0.7$ ms, $n = 33$) and nucleus ($\tau_{\mathrm{Diff}} = 1.6 \pm 0.6$ ms, $n = 29$). (**e**) Single-cell FRET efficiency histograms of ProTα in the cytosol and nucleus show a single population at $\langle E \rangle \approx 0.35$, similar to measurements in buffer. The shaded peak at $\langle E \rangle \approx 0$ originates from molecules lacking an active acceptor chromophore (**Supplementary Fig. 5**). (**f**) nsFCS curves obtained from averaging 51 cytosolic measurements yield information about the chain reconfiguration time of ProTα on the submicrosecond timescale (g_{DD}, donor-donor correlation; g_{AD}, acceptor-donor correlation; g_{AA}, acceptor-acceptor correlation).

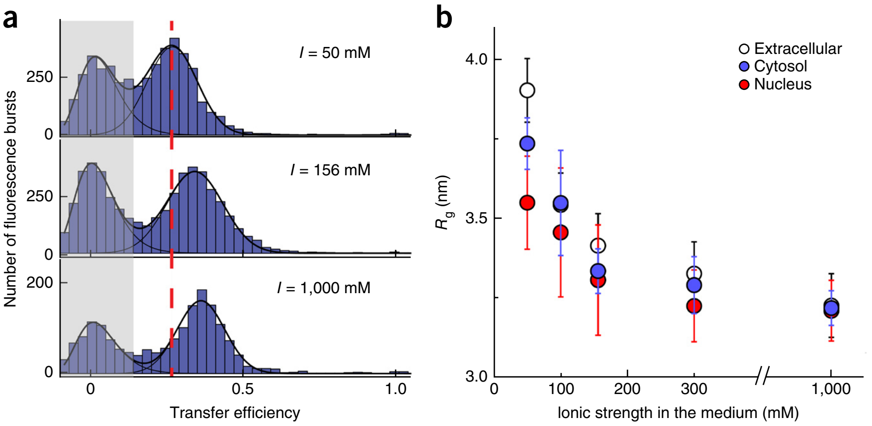


Figure 2 | Single-molecule FRET can detect small changes in protein conformation within live cells. (**a**) FRET efficiency histograms of ProTα in the cytosol of HeLa cells incubated in medium with different ionic strengths (the range of donor-only signal is shaded in gray). The red dashed line indicates the mean transfer efficiency of ProTα at I = 50 mM. (**b**) The collapse of ProTα (shown in terms of the radius of gyration, R_g, calculated from the transfer efficiencies; Online Methods) in the cytosol, nucleus and extracellular buffer. The s.d. of R_g for extracellular measurements was estimated from the variance of *in vitro* measurements performed on different instruments. The error bars for cytosolic and nuclear measurements represent the s.d. from cell-to-cell variation (Online Methods).

the emission of the injected sample from cellular autofluorescence (**Fig. 1b**), but imaging based on fluorescence intensity was adequate in most cases (**Supplementary Fig. 3**). We then acquired the actual single-molecule measurements by positioning the confocal volume at the cellular location of choice—for example, in the nucleus or the cytosol. If a fixed position was selected, predominantly freely diffusing molecules were observed; by slowly scanning across the region of interest, a defined volume could be probed (**Supplementary Fig. 4**) and potentially immobile molecules detected.

In contrast to measurements *in vitro*, the number of molecules available in microinjected cells is limited by the small cell volume. Measurements at the ~20–100 pM concentrations required for the confocal observation of single molecules were thus restricted to a few minutes before all molecules were photobleached. However, this time was sufficient for obtaining transfer efficiency histograms from single-cell measurements and, thus, assessing potential cell-to-cell variation of protein conformation (**Fig. 1e**). Compared to camera-based imaging approaches, whose strength is the parallel tracking of many molecules on longer timescales, confocal single-molecule detection provides additional observables from spectrally and polarization-selective single-photon counting[16], including fluorescence lifetimes (**Fig. 1b** and **Supplementary Fig. 5**), anisotropies (**Supplementary Fig. 6**) and, most importantly, dynamics in the millisecond range and even below[6,17,25–27].

Protein diffusivity, structure and conformational dynamics

We first tested the approach with confocal single-molecule measurements of the IDP[28,29] prothymosin α (ProTα)[28,30] in live cells. ProTα is a highly negatively charged protein that does not assume a folded structure under any known conditions and whose intramolecular distance distributions and dynamics have previously been characterized by single-molecule experiments *in vitro*[31–33]. First, fluorescence imaging after microinjection of ProTα labeled with Atto 532 and Biotium CF680R was used to monitor the intracellular distribution of the labeled protein (**Supplementary Fig. 2**). We quantified translational diffusion by correlating the fluorescence signal from the confocal measurements[6,26]. As expected, the intracellular diffusion time of ProTα increased relative to the diffusion time of the extracellular medium (**Fig. 1d**; to a similar extent in cytoplasm and nucleus), corresponding to an effective intracellular viscosity of 2.8 ± 1.1 mPa s (±s.d. from cell-to-cell variability unless otherwise specified; n = 39), which is within the range previously reported[34].

To quantify the dimensions of the protein, we used the distance dependence of intramolecular FRET. From measurements in individual cells, we obtained ~200–800 fluorescence bursts within the observation time of ~3 min before the majority of injected molecules were photobleached. We measured FRET efficiency histograms of ProTα with extracellular, cytosolic and nuclear localization (**Fig. 1e**). In all cases, a single population of FRET-labeled ProTα molecules with similar mean transfer efficiencies, $\langle E \rangle$, was observed. The cell-to-cell variation of $\langle E \rangle$ (s.d. = 0.02) was only slightly greater than its uncertainty from photon statistics (0.01) and close to the difference between cytosolic ($\langle E \rangle = 0.34 \pm 0.01$, $n = 8$) and nuclear measurements ($\langle E \rangle = 0.35 \pm 0.03$, n = 8). If we use a simple Gaussian chain model[32] to estimate the radius of gyration, R_g, of intracellular ProTα, we obtain 3.3 ± 0.1 nm (n = 8) in the cytosol and 3.3 ± 0.2 nm (n = 8) in the nucleus, very similar to the value in the extracellular medium (3.4 ± 0.1 nm, error propagated from the s.d. of $\langle E \rangle$) and to previous results *in vitro*[32]. The presence of an immobile subpopulation with different conformational properties was excluded by scanning over the intracellular region of interest during data acquisition (**Supplementary Fig. 4**).

For ProTα, as for other IDPs, nanosecond FCS (nsFCS) has been used to identify the characteristic chain-reconfiguration times on the 10- to 100-ns timescale[17,31]. We injected labeled ProTα and averaged the recordings from 51 cells with a total acquisition time of 2.5 h. The resulting submicrosecond correlation functions were of sufficient quality for us to discern the main features identified in previous *in vitro* measurements[17,25] (**Fig. 1f**): the drop of the correlation in the low nanosecond range owing to photon antibunching characteristic of single-molecule detection, as well as photon bunching with a slower decay and a time constant of 62 ± 6 ns (s.d. from splitting the data into three equal parts) originating from the intramolecular distance fluctuations in the IDP. These dynamics, which were slightly slower than those measured *in vitro* (24 ± 4 ns, n = 3 (15-h measurement each)), were in good agreement with the increased intracellular viscosity identified by translational diffusion measurements (**Fig. 1**). The presence of a broad and rapidly reconfiguring ensemble of conformations is further supported by the dependence of fluorescence lifetimes on transfer efficiency (**Supplementary Fig. 5**).

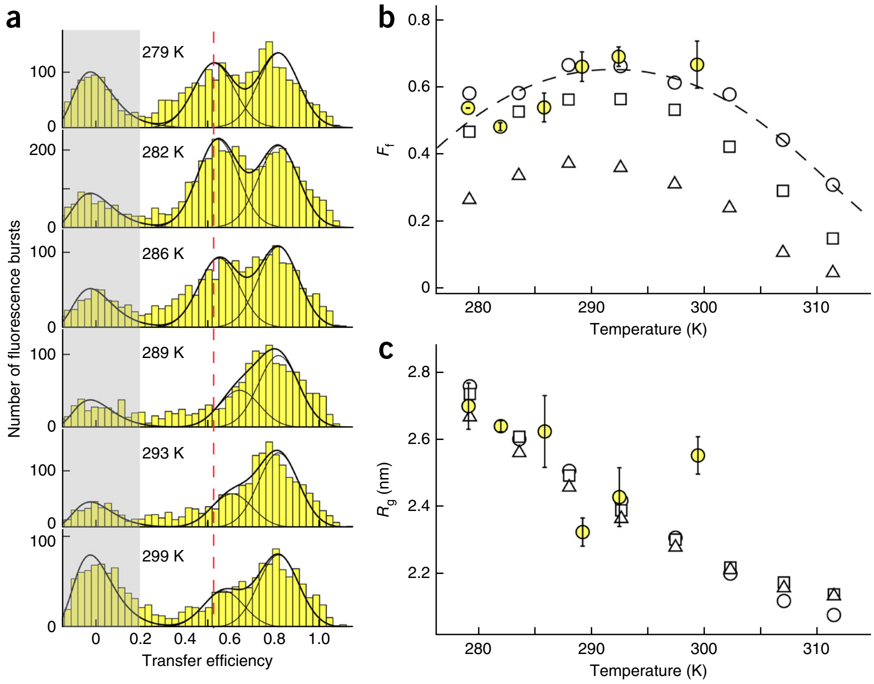


Figure 3 | Temperature dependence of protein stability and cold denaturation measured in live HeLa cells. (**a**) Single-molecule FRET efficiency histograms of Yfh1 microinjected into HeLa cells at different temperatures. The peak at $\langle E \rangle \approx 0.82$ corresponds to folded Yfh1; the peak at lower transfer efficiencies corresponds to unfolded Yfh1. For details on the fitting procedure, see Online Methods. (**b**) The fraction of folded Yfh1 (F_f) shows the temperature dependence of the intracellular conformational stability of Yfh1 (yellow circles), calculated as the ratio of the peak area of the folded state and the sum of the peak areas of folded and unfolded states (**a**). For comparison, the results of *in vitro* experiments in HBSS (white circles), HBSS without Ca^{2+} and Mg^{2+} (squares) and sodium phosphate, pH 7.0, (triangles) are shown. The intracellular temperature dependence of F_f was fitted assuming a constant heat capacity change of unfolding (dashed line), as previously described[40]. (**c**) The temperature dependence of the radius of gyration (R_g; Online Methods) of unfolded Yfh1 inside HeLa cells (yellow circles) and in buffer (HBSS, white circles; HBSS without Ca^{2+} and Mg^{2+}, squares; sodium phosphate, pH 7.0, triangles). The data points of the in-cell measurements in **b**,**c** were obtained from the global fit shown in **a**. Error bars correspond to s.d. estimated by splitting each data set into two and treating F_f or R_g of the unfolded population determined from each half as independent measurements.

The notable similarity of the intra- and extracellular behavior of ProTα may raise the question of whether in-cell single-molecule FRET experiments are sensitive enough to detect small changes in conformation. The intramolecular distance distributions of the highly negatively charged ProTα have previously been shown to be very responsive to charge screening with changes in ionic strength[32,33]. We mimicked this situation by increasing the extracellular salt concentration, which is expected to lead to a transient increase in intracellular ionic strength[35]. Indeed, the resulting decrease in the R_g of intracellular ProTα as derived from the transfer efficiencies showed a similar trend to the extracellular behavior (**Fig. 2**), indicating that even small changes in intramolecular distances of 0.2–0.3 nm were detectable in cells.

Temperature-dependent intracellular protein stability

Detecting changes in conformational stability *in vivo* is another important aspect for identifying the effect of cellular factors on biomolecular behavior[3,36,37]. We investigated the effect of temperature on the conformational properties of the yeast homolog of frataxin (Yfh1)[38–40]. Frataxin exhibits cold denaturation in a readily accessible temperature range, i.e., it unfolds not only at high temperature but also below ~290 K. We labeled the protein with Atto 532 and Abberior Star 635 and developed a temperature-controlled stage optimized for cell culture observation and microinjection combined with single-molecule detection (**Supplementary Fig. 7**).

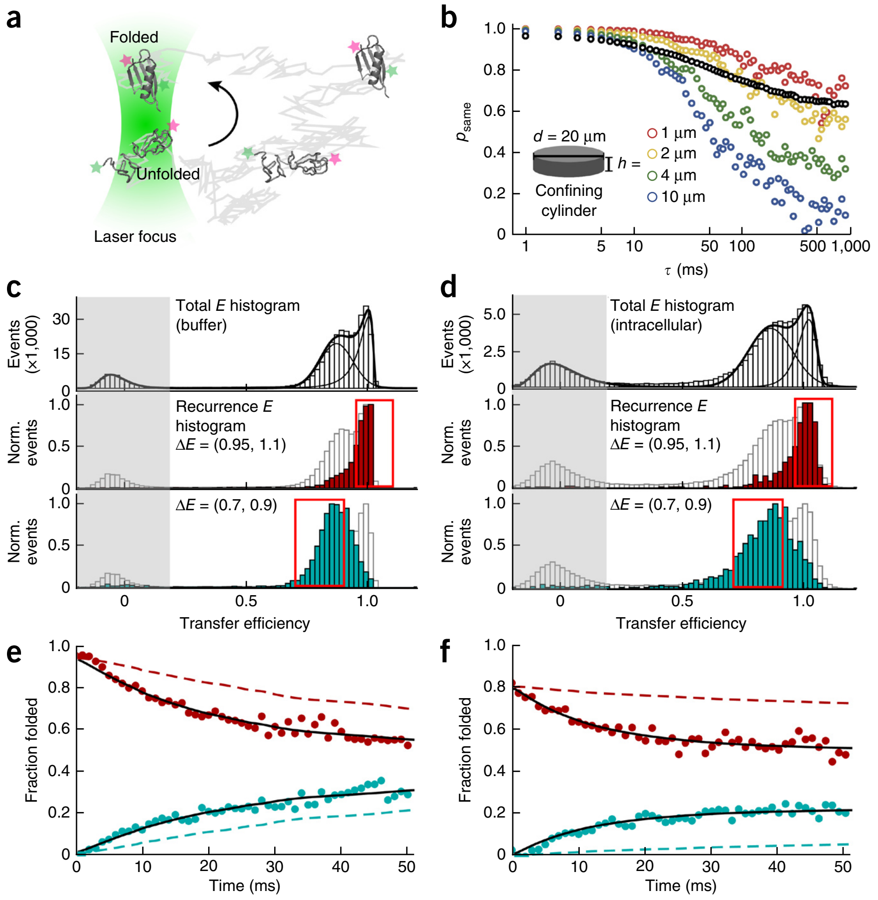
Transfer efficiency histograms of labeled frataxin after its injection into live cells at different temperatures showed two populations with a folded state at $\langle E \rangle \approx 0.82$ and an unfolded state at lower transfer efficiency (**Fig. 3**). In spite of the lower signal-to-noise ratio, the data reflected the behavior observed *in vitro*[40] (**Supplementary Fig. 8**). By extracting the fraction of folded protein from the peak areas, we detected a maximum in conformational stability of the protein at ~290 K and a decrease at both lower and higher temperatures, corresponding to cold and heat denaturation. The behavior was very sensitive to buffer conditions, especially to salt concentrations[39] (**Fig. 3b**). In HBSS buffer, the stability was at the same level as in cells, pointing toward good consistency of this optimized cell culture medium with the cellular ion concentrations. Measurements in HBSS buffer without Ca^{2+} and Mg^{2+} confirmed the pronounced sensitivity of frataxin stability to divalent ions[39]. The intracellular single-molecule measurements of frataxin were consistent with the compaction of its unfolded state with increasing temperature previously observed in buffer[40] (**Fig. 3c**).

Single-molecule protein-folding kinetics in the cell

Single-molecule spectroscopy enables kinetics to be extracted from equilibrium measurements, a quality that is particularly valuable in the cell, where the choice of perturbations compatible with cellular homeostasis is very limited. As a model for metastable cellular proteins that exhibit equilibrium conformational interconversion dynamics, we investigated a variant of GB1, the IgG-binding domain of protein G labeled with Atto 532 and Atto 647N. Its conformational stability and folding kinetics in the millisecond range have been well characterized *in vitro*[41] and *in vivo*[42]. By using a slightly destabilized variant (Online Methods), we could populate both folded and unfolded states at equilibrium under physiological conditions, as reflected by the two peaks in the transfer efficiency histogram at $\langle E \rangle = 1.00$ and 0.85, respectively (**Fig. 4** and **Supplementary Fig. 9**).

To extract the rate of interconversion between the folded and unfolded conformations at equilibrium, we used recurrence analysis of single particles (RASP), an approach that has recently

Figure 4 | Intracellular protein-folding dynamics from recurrence analysis (RASP)[43]. (**a**) Schematic of a protein recurring to the confocal volume; in this example, the protein happens to fold between detected fluorescence bursts. (**b**) Comparison of the experimental p_{same} curve obtained in cells (black) with simulated p_{same} curves of proteins in a confining cylinder (diameter, 20 µm; height, 1–10 µm; protein concentration, 100 pM). (**c**,**d**) FRET efficiency histograms of GB1 in 50 mM phosphate buffer (**c**) and inside cells (**d**). The range of donor-only signal is shaded in gray, and the initial transfer efficiency range, ΔE, is indicated with red boxes. Recurrence FRET efficiency histograms show the folded (red, $\Delta E = (0.95–1.1)$) and unfolded (cyan, $\Delta E = (0.7–0.9)$) populations of GB1 within a recurrence interval $T = (0, 1\ \mathrm{ms})$ (measurements performed at 4 °C; Online Methods). (**e**,**f**) Fractions of folded molecules from recurrence FRET efficiency histograms with increasing recurrence intervals (**Supplementary Fig. 10**) starting from unfolded (cyan) and folded molecules (red). Global fits (solid lines) yielded relaxation times of 21 ± 4 ms (buffer, **e**) and 12 ± 2 ms (intracellular, **f**). Dashed lines show the behavior expected in the absence of folding dynamics (i.e., from the appearance of new molecules only), as calculated from burst time correlation analysis[43]. Differences between dashed and solid lines indicate population interconversion on the millisecond timescale (for a control measurement with noninterconverting species, see **Supplementary Fig. 11**).



been shown *in vitro* to allow quantification of conformational changes on timescales from ~50 µs to ~100 ms (ref. 43; **Fig. 4**, Online Methods and **Supplementary Fig. 10**). Cellular confinement strongly increases the same-molecule probability, $p_{\mathrm{same}}(\tau)$, i.e., the probability that two fluorescence bursts separated by time τ are emitted by the same molecule, as illustrated by simulations in a cylindrical volume (**Fig. 4b**). For cylinder heights of 1–2 µm (dimensions similar to those of an adherent HeLa cell), the p_{same} decays closely resemble intracellular measurements, with p_{same} values of around 60% for timescales of seconds. This effect extends the effective observation time and makes recurrence analysis ideally suited for probing intracellular conformational dynamics.

RASP yielded a relaxation time for GB1 folding-unfolding of 21 ± 4 ms in buffer (error from the fit, valid for all decay times shown here; equilibrium constant $K_{\mathrm{eq}} = 1.3 \pm 0.1$), in agreement with the time of 23 ± 7 ms ($K_{\mathrm{eq}} = 1.5 \pm 0.1$) obtained from single-molecule perturbation experiments using microfluidic mixing under the same solution conditions (Online Methods and **Supplementary Fig. 9**) and in the range expected from previous work[41]. In HeLa cells, the transfer efficiency histogram of the microinjected GB1 variant was broadened compared to buffer measurements—for example, from the additional background contribution; nevertheless, two populations could be identified, as further supported by the recurrence FRET efficiency histograms, which allow the peaks from folded and unfolded protein to be clearly separated[43] (**Fig. 4** and **Supplementary Fig. 10**). We observed, on the basis of the peak integrals, a similar fraction of folded state (0.33 ± 0.03) to that in buffer (0.43 ± 0.02). RASP analysis of the intracellular measurements yielded a folding-unfolding relaxation time of 12 ± 2 ms ($K_{\mathrm{eq}} = 2.0 \pm 0.2$; **Fig. 4**), indicating at most a small effect of the cellular environment—such as from macromolecular crowding—on the conformational stability and folding dynamics of GB1 (refs. 37,44). We demonstrated the robustness of the RASP method by performing control measurements with noninterconverting species in buffer and injected into cells (**Supplementary Fig. 11**).

DISCUSSION

The results presented here show that an integrated approach addressing the multifactorial challenges of in-cell single-molecule spectroscopy enables the application of a broad spectrum of confocal single-molecule fluorescence methods in eukaryotic cells. With single-molecule FRET and the resulting separation of subpopulations, not only can intramolecular distances be quantified, but the conformational stability of proteins can be measured in live cells and compared to *in vitro* observations (**Figs. 3** and **4**), including the effects of temperature. The dimensions of intrinsically disordered and unfolded proteins are accessible (**Figs. 1–3**), as are their responses to changes in intracellular conditions, such as ion concentrations and temperature (**Figs. 2** and **3**). Most notably, the approach enables in-cell conformational dynamics to be investigated, illustrated here by the use of nsFCS for quantifying submicrosecond reconfiguration times (**Fig. 1**) and the analysis of FRET efficiency histograms with recurrence analysis to extract millisecond protein-folding kinetics (**Fig. 4**). Together with the well-known possibility of determining intracellular diffusivity with FCS[6,26], the effective intracellular viscosity can be monitored from molecular to cellular length scales. All of these observables are accessible as a function of cellular localization (**Fig. 1**) and thus enable the investigation of spatial variations in conformation and dynamics. The availability of FLIM can aid the discrimination between the signal of the injected sample and autofluorescence (**Fig. 1**).

Our observations exemplify how this methodology can provide new possibilities for probing the properties of biomolecules in live cells. The single-molecule FRET experiments on ProTα clearly showed that this IDP remains unstructured in the cell, as judged from its dimensions and submicrosecond chain dynamics (**Fig. 1**). Similar conclusions were previously drawn for α-synuclein via intracellular NMR-based experiments[45,46], indicating that even though many IDPs are known to form a folded structure upon interaction with their cellular targets[29], IDPs can stay unfolded upon binding and in live cells. The results on protein GB1 and frataxin indicated a remarkable robustness in the conformational stabilities and even folding kinetics of these proteins in cells compared to in simple buffered solutions (**Figs. 3** and **4**). However, even modest changes in quantitative behavior may have important functional or regulatory consequences[2,36]. The sensitivity of the in-cell single-molecule approach (**Figs. 2** and **3**) is thus expected to be ideally suited for studying the effects of molecular crowding[44] and the cellular machinery on the conformations, dynamics and intracellular interactions of biological macromolecules, even as a function of cellular localization.

Our approach complements existing single-molecule methods in eukaryotes[14] and prokaryotes[10–13] and other techniques. In-cell NMR spectroscopy, for example, can provide atomistic detail and a wide range of dynamics[37,42,45,47], but it is currently limited to averages over large molecular ensembles and many cells. Recent advances in quantitative cellular FRET imaging based on fusions with fluorescent proteins and temperature jumps have led to the spatiotemporal mapping of protein stability and kinetics with high sensitivity[3,36], but this ensemble technique requires rapid changes in temperature as a perturbation and is currently limited to timescales in the seconds range and above. Finally, super-resolution microscopy[8,9] now routinely enables fluorescence imaging with spatial resolution of tens of nanometers, but conformational changes and dynamics at the time and length scales investigated here have remained beyond its reach. We therefore expect that single-molecule spectroscopy will play an important role in bridging the gap between our quantitative understanding of biomolecules *in vitro* and *in vivo*.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank S. Weidner and A. Schmid for excellent technical assistance in instrumentation development, A. Holla and H. Hofmann for discussions regarding the choice of fluorophores, A. Flütsch (University of Zurich) for providing HeLa cells, and K. Buholzer for assistance with culturing cells. B.S. was supported by the Swiss National Science Foundation and the European Research Council. A.Z.-A. was supported by the Human Frontier Science Program. M.A. was supported by a Postdoctoral Fellowship from the Forschungskredit of the University of Zurich (FK-13-034). J.C.S. was supported by a Chemiefonds fellowship of the German Chemical Industry Fund. Differential interference contrast imaging was performed at the Center for Microscopy and Image Analysis, University of Zurich.

AUTHOR CONTRIBUTIONS

B.S. designed the project together with I.K., A.Z.-A. and M.A.; I.K., A.Z.-A. and M.A. performed the experiments; I.K., A.Z.-A., M.A., A.S. and D.N. analyzed the data; J.C.S. and I.K. conducted cell culture experiments; A.Z.-A., B.W. and F.D. performed and analyzed microfluidics experiments; I.K., M.A., B.W. and D.N. developed instrumentation; A.P., A.S., D.N. and B.S. supervised the project; I.K., A.Z.-A., M.A., D.N. and B.S. wrote the paper with the help of all authors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Shaner, N.C., Steinbach, P.A. & Tsien, R.Y. A guide to choosing fluorescent proteins. *Nat. Methods* **2**, 905–909 (2005).
2. Phillip, Y., Kiss, V. & Schreiber, G. Protein-binding dynamics imaged in a living cell. *Proc. Natl. Acad. Sci. USA* **109**, 1461–1466 (2012).
3. Ebbinghaus, S., Dhar, A., McDonald, J.D. & Gruebele, M. Protein folding stability and dynamics imaged in a living cell. *Nat. Methods* **7**, 319–323 (2010).
4. Fenz, S.F., Pezzarossa, A. & Schmidt, T. in *Comprehensive Biophysics* Vol. 2 (ed. Egelman, E.H.) Ch. 13, 260–272 (Elsevier, 2012).
5. Pitchiaya, S., Heinicke, L.A., Custer, T.C. & Walter, N.G. Single molecule fluorescence approaches shed light on intracellular RNAs. *Chem. Rev.* **114**, 3224–3265 (2014).
6. Mütze, J., Ohrt, T. & Schwille, P. Fluorescence correlation spectroscopy *in vivo*. *Laser Photon. Rev.* **5**, 52–67 (2011).
7. Digman, M.A. & Gratton, E. Scanning image correlation spectroscopy. *Bioessays* **34**, 377–385 (2012).
8. Hell, S.W. Far-field optical nanoscopy. *Science* **316**, 1153–1158 (2007).
9. Jones, S.A., Shim, S.H., He, J. & Zhuang, X. Fast, three-dimensional super-resolution imaging of live cells. *Nat. Methods* **8**, 499–508 (2011).
10. Persson, F., Barkefors, I. & Elf, J. Single molecule methods with applications in living cells. *Curr. Opin. Biotechnol.* **24**, 737–744 (2013).
11. Stratmann, S.A. & van Oijen, A.M. DNA replication at the single-molecule level. *Chem. Soc. Rev.* **43**, 1201–1220 (2014).
12. Uphoff, S., Reyes-Lamothe, R., Garza de Leon, F., Sherratt, D.J. & Kapanidis, A.N. Single-molecule DNA repair in live bacteria. *Proc. Natl. Acad. Sci. USA* **110**, 8063–8068 (2013).
13. Crawford, R. *et al.* Long-lived intracellular single-molecule fluorescence using electroporated molecules. *Biophys. J.* **105**, 2439–2450 (2013).
14. Sakon, J.J. & Weninger, K.R. Detecting the conformation of individual proteins in live cells. *Nat. Methods* **7**, 203–205 (2010).
15. Selvin, P.R. & Ha, T. *Single-Molecule Techniques: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 2008).
16. Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P.J. & Seidel, C.A.M. Accurate single-molecule FRET studies using multiparameter fluorescence detection. *Methods Enzymol.* **475**, 455–514 (2010).
17. Schuler, B. & Hofmann, H. Single-molecule spectroscopy of protein folding dynamics—expanding scope and timescales. *Curr. Opin. Biotechnol.* **23**, 36–47 (2013).
18. Banerjee, P.R. & Deniz, A.A. Shedding light on protein folding landscapes by single-molecule fluorescence. *Chem. Soc. Rev.* **43**, 1172–1188 (2014).
19. Verdurmen, W.P.R., Luginbühl, M., Honegger, A. & Plückthun, A. Efficient cell-specific uptake of binding proteins into the cytoplasm through engineered modular transport systems. *J. Control. Release* **200**, 13–22 (2015).
20. Aubin, J.E. Autofluorescence of viable cultured mammalian cells. *J. Histochem. Cytochem.* **27**, 36–43 (1979).
21. McNeil, P.L., Murphy, R.F., Lanni, F. & Taylor, D.L. A method for incorporating macromolecules into adherent cells. *J. Cell Biol.* **98**, 1556–1564 (1984).
22. Morgan, W.F. & Day, J.P. The introduction of proteins into mammalian cells by electroporation. *Methods Mol. Biol.* **48**, 63–71 (1995).
23. Milletti, F. Cell-penetrating peptides: classes, origin, and current landscape. *Drug Discov. Today* **17**, 850–860 (2012).
24. Berezin, M.Y. & Achilefu, S. Fluorescence lifetime measurements and biological imaging. *Chem. Rev.* **110**, 2641–2684 (2010).

25. Nettels, D., Gopich, I.V., Hoffmann, A. & Schuler, B. Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc. Natl. Acad. Sci. USA* **104**, 2655–2660 (2007).
26. Rigler, R. & Elson, E.S. *Fluorescence Correlation Spectroscopy: Theory and Applications* (Springer, 2001).
27. Gopich, I.V. & Szabo, A. in *Single-Molecule Biophysics: Experiment and Theory* Vol. 146 (eds. Komatsuzaki, T., Kawakami, M., Takahashi, S., Yang, H. & Silbey, R.J.) 245–297 (Wiley, 2011).
28. Uversky, V.N. What does it mean to be natively unfolded? *Eur. J. Biochem.* **269**, 2–12 (2002).
29. Wright, P.E. & Dyson, H.J. Linking folding and binding. *Curr. Opin. Struct. Biol.* **19**, 31–38 (2009).
30. Gast, K. *et al.* Prothymosin alpha: a biologically active protein with random coil conformation. *Biochemistry* **34**, 13211–13218 (1995).
31. Soranno, A. *et al.* Quantifying internal friction in unfolded and intrinsically disordered proteins with single molecule spectroscopy. *Proc. Natl. Acad. Sci. USA* **109**, 17800–17806 (2012).
32. Müller-Späth, S. *et al.* Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. USA* **107**, 14609–14614 (2010).
33. Soranno, A. *et al.* Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments. *Proc. Natl. Acad. Sci. USA* **111**, 4874–4879 (2014).
34. Kalwarczyk, T. *et al.* Comparative analysis of viscosity of complex liquids and cytoplasm of mammalian cells at the nanoscale. *Nano Lett.* **11**, 2157–2163 (2011).
35. Burg, M.B., Ferraris, J.D. & Dmitrieva, N.I. Cellular response to hyperosmotic stresses. *Physiol. Rev.* **87**, 1441–1474 (2007).
36. Guo, M., Xu, Y. & Gruebele, M. Temperature dependence of protein folding kinetics in living cells. *Proc. Natl. Acad. Sci. USA* **109**, 17863–17867 (2012).
37. Smith, A.E., Zhang, Z., Pielak, G.J. & Li, C. NMR studies of protein folding and binding in cells and cell-like environments. *Curr. Opin. Struct. Biol.* **30**, 7–16 (2015).
38. Pastore, A. *et al.* Unbiased cold denaturation: low- and high-temperature unfolding of yeast frataxin under physiological conditions. *J. Am. Chem. Soc.* **129**, 5374–5375 (2007).
39. Sanfelice, D. *et al.* Yeast frataxin is stabilized by low salt concentrations: cold denaturation disentangles ionic strength effects from specific interactions. *PLoS ONE* **9**, e95801 (2014).
40. Aznauryan, M., Nettels, D., Holla, A., Hofmann, H. & Schuler, B. Single-molecule spectroscopy of cold denaturation and the temperature-induced collapse of unfolded proteins. *J. Am. Chem. Soc.* **135**, 14040–14043 (2013).
41. Kuszewski, J., Clore, G.M. & Gronenborn, A.M. Fast folding of a prototypic polypeptide: the immunoglobulin binding domain of streptococcal protein G. *Protein Sci.* **3**, 1945–1952 (1994).
42. Selenko, P., Serber, Z., Gadea, B., Ruderman, J. & Wagner, G. Quantitative NMR analysis of the protein G B1 domain in *Xenopus laevis* egg extracts and intact oocytes. *Proc. Natl. Acad. Sci. USA* **103**, 11904–11909 (2006).
43. Hoffmann, A. *et al.* Quantifying heterogeneity and conformational dynamics from single molecule FRET of diffusing molecules: recurrence analysis of single particles (RASP). *Phys. Chem. Chem. Phys.* **13**, 1857–1871 (2011).
44. Zhou, H.-X., Rivas, G. & Minton, A.P. Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.* **37**, 375–397 (2008).
45. Binolfi, A., Theillet, F.X. & Selenko, P. Bacterial in-cell NMR of human α-synuclein: a disordered monomer by nature? *Biochem. Soc. Trans.* **40**, 950–954 (2012).
46. Waudby, C.A. *et al.* In-cell NMR characterization of the secondary structure populations of a disordered conformation of α-synuclein within *E. coli* cells. *PLoS ONE* **8**, e72286 (2013).
47. Tochio, H. Watching protein structure at work in living cells using NMR spectroscopy. *Curr. Opin. Chem. Biol.* **16**, 609–613 (2012).

ONLINE METHODS

Protein expression and labeling. Fluorescent dyes for single-molecule experiments were chosen according to their availability as a maleimide derivative for site-specific labeling and their photophysical (photostability and brightness) and chemical (hydrophobicity) properties. For probing conformational dynamics on the length and time scales required here, genetically introduced fusions with fluorescent proteins are of limited use owing to their relatively large size, lower brightness, and photostability[48]; the complication of heterogeneous expression levels[49]; and the tendency of forming oligomers[50]. We thus chose to work with organic fluorophores. The broad availability of fluorophores with absorption maxima in the vicinity of the popular 532-nm laser line facilitates the choice of a FRET donor and a corresponding acceptor dye. As a donor dye, we used Atto 532 maleimide (ATTO-TEC GmbH). For acceptor dyes, we covered a range from hydrophilic to hydrophobic dyes, specifically Biotium CF680R maleimide (Biotium Inc.), Abberior Star 635 maleimide (Abberior GmbH) and Atto 647N maleimide (ATTO-TEC GmbH). The Förster radii for the dye-pairs were calculated to be 5.7 nm for Atto 532/Biotium CF680R, 6.0 nm for Atto 532/Abberior Star 635, and 6.0 nm for Atto 532/Atto 647N on the basis of the quantum yields and absorption and emission spectra provided by the manufacturer and using an averaged orientation factor of $\kappa^2 = 2/3$, as supported by single-molecule anisotropy measurements *in vitro* and *in vivo* (**Supplementary Fig. 6** and Online Methods section "Data analysis").

The three different acceptor fluorophores (Atto 647N, Abberior Star 635 and Biotium CF680R), in combination with the donor dye Atto 532, were used to assess the robustness of different FRET pairs for measurements *in vitro* and *in vivo*. The selected acceptor dyes have absorption maxima between 634 and 680 nm and emission maxima between 654 and 701 nm (as specified by the manufacturer), in a suitable range to act as acceptors for FRET from Atto 532. Atto 647N is particularly popular for experiments both *in vitro* and *in vivo* owing to its very high photostability and brightness[51]. Abberior Star 635 was recently developed for demanding applications such as stimulated emission depletion microscopy[52]. Biotium CF680R was chosen on the basis of a screen we performed for dyes with suitably high photostability and brightness combined with low hydrophobicity; we are not aware of any previous use in single-molecule spectroscopy. A major difference between the dyes is their hydrophobicity, with Biotium CF680R being the most hydrophilic, Abberior Star 635 being moderately hydrophobic, and Atto 647N exhibiting the highest hydrophobicity according to the elution times in reversed-phase HPLC. Our results showed that the FRET pairs Atto 532/Biotium CF680R on ProTα (**Figs. 1** and **2**), Atto532/Abberior Star 635 on frataxin (**Fig. 3**) and Atto 532/Atto 647N on GB1 (**Fig. 4**) are all suitable for in-cell single-molecule spectroscopy, indicating that the method is not limited to specific fluorophores. However, care has to be taken regarding the hydrophobicity of the dyes. Especially Atto647N is known for its hydrophobicity[53], which can lead to nonspecific interactions with other molecules or even the labeled protein itself. It is therefore essential to characterize labeled proteins *in vitro* before using them in *in vivo* experiments: for example, regarding the conformational stability (**Fig. 3**), the fluorescence lifetime of the attached dyes (**Supplementary Fig. 5**), or fluorescence anisotropy (**Supplementary Fig. 6**). In the cases of ProTα, Frataxin and GB1, no interfering interactions were observed after we labeled them with the dye pairs used here.

Prothymosin α (ProTα). Cysteine residues for the specific labeling of ProTα using maleimide chemistry were introduced by site-directed mutagenesis at the positions 1 and 56 (**Supplementary Table 1**). ProTα was produced in *Escherichia coli* BL21 with an N-terminal His-tag for purification. Cells were grown in LB medium, expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG, Sigma), and the harvested cells were disrupted. As previously described[33], ProTα was purified with a Ni-NTA column (Thermo Scientific) after ammonium sulfate precipitation[54]. The His-tag was cleaved off with HRV 3C protease (recombinantly produced in-house as a hexahistidine fusion), followed by a second Ni-NTA chromatography run to remove the His-tag and the protease. ProTα was reduced with 10 mM tris(2-carboxyethyl) phosphine (TCEP, Sigma) and further purified by reversed-phase high-pressure liquid chromatography (RP-HPLC) on a Reprosil Gold 200 column (Dr. Maisch GmbH) with an acetonitrile gradient. Purified ProTα was lyophilized in a SpeedVac concentrator (Thermo Scientific), dissolved in 0.1 M sodium phosphate (Sigma) buffer, pH 7.2, and labeled with Atto 532 maleimide according to the supplier's manual at a molar ratio of dye to protein of 0.8:1. The reaction was stopped with β-mercaptoethanol (Sigma), reduced with 10 mM TCEP and purified by RP-HPLC on an XTerra C18 column (Waters), which enabled not only the separation of singly labeled, doubly labeled and unlabeled protein but also the separation of the singly labeled permutants, as confirmed by proteolytic digest and electrospray ionization mass spectrometry (ESI-MS). The fraction containing one of the singly labeled ProTα permutants was lyophilized in a SpeedVac concentrator and was labeled and purified analogously with a molar excess of Biotium CF680R maleimide. The mass of the doubly labeled ProTα was confirmed by ESI-MS.

Yeast frataxin homolog Yfh1. The expression and purification of Yfh1 (**Supplementary Table 1**) was performed as previously described[40]. Purified Yfh1 was first labeled with Atto 532 maleimide in 6 M guanidinium chloride (GdmCl), 100 mM sodium phosphate, pH 7.1, at a molar ratio of dye to protein of 0.7:1 for 2 h at room temperature. Singly labeled Yfh1 was isolated from the reaction mixture by RP-HPLC using a Reprosil Gold C18 column with a water-acetonitrile gradient, which enabled not only the separation of singly labeled, doubly labeled and unlabeled protein but also the separation of the singly labeled permutants. The lyophilized fractions containing one of the singly labeled permutants of Yfh1 were further labeled with Abberior Star 635 maleimide. A twofold molar excess of Abberior Star 635 was reacted with singly Atto 532–labeled protein in 6 M GdmCl, 100 mM sodium phosphate, pH 7.1, for 2 h at room temperature. The purification of the doubly labeled Yfh1 was again performed by RP-HPLC, as described above. The mass of doubly labeled Yfh1 was confirmed by ESI-MS.

IgG-binding domain of protein G (GB1). A variant of GB1 (ref. 55) containing the destabilizing F30L-G41A substitutions[56] (**Supplementary Table 1**) was site-specifically labeled at positions Q2C and T55C (very close to the termini to achieve the lowest possible transfer efficiency compared to the folded state) with the maleimide derivatives of the donor Atto 532 and the acceptor Atto 647N, respectively. GB1 was expressed with an N-terminal His-tag followed by a cleavage site for tobacco etch

doi:10.1038/nmeth.3475

virus (TEV) protease using *E. coli* Rosetta DE3 cells (Millipore). The cells were cotransfected with the pRARE plasmid (Millipore) to further boost the expression of GB1. The Rosetta cells were grown in 2YT medium, and expression was induced with IPTG. The harvested cells were lysed in 6 M GdmCl (Sigma), 100 mM sodium phosphate, 10 mM Tris-HCl (Sigma), 10 mM imidazole (Sigma), pH 8.0 (binding buffer), by rocking them at room temperature for 1 h. Insoluble proteins and cellular debris were then pelleted, and the supernatant was incubated with 5 ml Ni-NTA agarose slurry (Thermo Scientific). The slurry was gently pelleted and washed with binding buffer four times. Bound GB1 was eluted by incubating the slurry in 6 M GdmCl, 0.2 M acetic acid (Sigma), pH 4.5, for 20 min, rocking at room temperature. The supernatant containing GB1 was filtered through a 0.45-µm filter (TPP) and refolded by equilibrium dialysis (3-kDa-cutoff membrane, Spectrum Laboratories) against three times 4 liters of 50 mM sodium phosphate, 300 mM sodium chloride (Sigma), pH 7.0, at 4 °C. TEV protease (recombinantly produced in-house as a hexahistidine fusion) was added to GB1 at a molar ratio of 50:1 (GB1/TEV) along with 5 mM dithiothreitol (Sigma) to cleave off the His-tag, followed by dialysis against 50 mM sodium phosphate, 300 mM NaCl, pH 8.0, as above. GB1 was then subjected to a second round of Ni-NTA pulldown (equilibrated in the same buffer) to remove the cleaved His-tag as well as the TEV protease. GB1 was further purified with size-exclusion chromatography on a Superdex 75 column (GE Healthcare) in 50 mM potassium phosphate, 100 mM NaCl, 0.5 mM EDTA (Sigma), at pH 7.5. Fractions containing GB1 were collected and were subjected to RP-HPLC on a Reprosil Gold C18 column to remove additional degradation products. Protein samples were reduced with TCEP and eluted with an acetonitrile gradient. Samples were subsequently lyophilized and labeled with Atto 532 maleimide at a molar ratio of dye to protein of 0.7:1. The reaction products were purified by RP-HPLC using an acetonitrile gradient on an XTerra C18 column, which enabled not only the separation of singly labeled, doubly labeled and unlabeled protein but also the separation of the singly labeled permutants. The purified protein fraction containing one of the singly labeled permutants was subsequently lyophilized and labeled with Atto 647N maleimide with a molar ratio of protein to dye of 1:3. The reaction mixture was purified by reversed-phase HPLC using an acetonitrile gradient on an XTerra C18 column, as above. The mass of doubly labeled GB1 was confirmed by ESI-MS. Yields of correctly FRET-labeled samples were >90% for all three proteins, as indicated by HPLC elution profiles, mass spectrometry, and the low 'donor-only' signal in single-molecule measurements *in vitro*.

Cell culture and microinjection. *Cell culture.* HeLa cells (American Type Culture Collection; mycoplasma free, tested with the MycoAlert mycoplasma detection Kit (Lonza); the identity of the cells was confirmed by visually inspecting the cell morphology) were cultured in tissue culture flasks (75 cm^2, TPP) in humid atmosphere at 37 °C and 5% CO_2. The cells were grown in phenol red–free Dulbecco's modified Eagle medium (DMEM, Invitrogen), supplemented with 10% FCS (Bioconcept Amimed), 100 units/ml penicillin, and 0.1 mg/ml streptomycin (Sigma). Cells were subcultured when approximately 80% confluency was reached using trypsin-EDTA (Sigma) for cell detachment. One day before a single-molecule experiment, 2×10^5 HeLa cells in 4 ml phenol red–free DMEM medium were pipetted into a microscopy dish (35-mm glass-bottom dishes, Ibidi) and incubated overnight (37 °C, 5% CO_2) until the cells became adherent again. Prior to single-molecule measurements, the medium was aspirated; the cells were then washed three times with Hanks balanced salt solution (HBSS, Sigma; HBSS contains 1.3 mM $CaCl_2$, 0.8 mM $MgSO_4$, 5.4 mM KCl, 0.44 mM KH_2PO_4, 4.2 mM $NaHCO_3$, 136.9 mM NaCl, 0.3 mM Na_2HPO_4 and 5.6 mM glucose) and covered with a buffer of choice. For experiments with different salt concentrations (up to 300 mM total ionic strength in the medium) or for temperature experiments, the microscopy dishes were first coated with poly(L-lysine) solution (0.01% in H_2O, Sigma) to improve cell adhesion.

Microinjection. Protein solutions were injected into HeLa cells using a Xenoworks microinjection system (Sutter Instrument), which was directly attached to the single-molecule fluorescence microscope. The pressure pulses (0.5–1 bar, 20–50 ms) for the injections were delivered by a home-built injector or a Picospritzer III (Parker). The sample was injected into cells using Femtotips II capillaries (Eppendorf). Before cell injection, the operational capability of the system was tested by injecting sample several times into the cell medium while recording a transfer efficiency histogram with the single-molecule fluorescence microscope. For all experiments with GB1 and Yfh1, the inside of the capillaries was first coated with 1 mg/ml poly(L-lysine)-*graft*-poly(ethylene glycol)[57] (PLL(20)-*g*[3.5]-PEG(2), SuSoS), dissolved in doubly distilled H_2O (ddH_2O), for 5 min, to reduce the adherence of the proteins to the glass capillary. Before loading the sample, which needs to be dissolved in a low-salt buffer to minimize dissociation of PLL-PEG—for example, 20 mM sodium phosphate buffer, pH 7.0—the coated capillaries were rinsed twice with ddH_2O. All protein samples for injections contained 0.005% Tween 20 (Pierce) to reduce adsorption of protein to any container surfaces used during dilution, and the sample was injected into either the cytosol or the nucleus of a targeted cell. Solutions for injection contained labeled protein at concentrations of ~20–50 nM. For every measurement, a single pressure pulse was applied to deliver the sample into the cell. The pressure settings and the duration of the pulses for the sample injections into HeLa cells were optimized in a way that (i) enough sample was injected to observe a sufficiently large number of fluorescence bursts per cell and (ii) the pressure pulse was gentle enough to avoid the formation of membrane protuberances or cell lysis, clear indications of compromised cellular viability (**Supplementary Fig. 12**). Slight adjustments of pressure pulses within the range given above were made for every set of measurements involving a stock solution of a given protein concentration. Longer-term cell viability after injection with parameters optimized in this way was assessed by adding propidium iodide[58] (Sigma) to the medium after injection of buffer into the cytosol of 23 HeLa cells (**Supplementary Fig. 1**). The membrane integrity and cell morphology were tracked up to 1.5 h after injection. Immediately after injection, all injected cells were still intact. After 1.5 h, 74% of all injected cells were still intact; four cells (18%) showed blebbing, and one cell was found dead according to the propidium iodide stain.

Instrumentation for single-molecule FRET in cells. *Single-molecule fluorescence microscope.* Single-molecule fluorescence measurements were performed on a custom-built confocal

doi:10.1038/nmeth.3475

instrument using an Olympus IX71 microscope body. Light from a 20-MHz supercontinuum laser (SC-450-4, Fianium) was used for excitation, with the wavelength selected via a 520/15 band-pass filter (Chroma Technology), and focused into the sample with an Olympus UplanApo 60×/1.20-W objective (Olympus). Photons emitted from the sample were collected through the same objective. Scattered excitation light was eliminated with a long-pass filter (LP532, Chroma Technology) before the emitted photons passed a 100-μm confocal pinhole. The emitted photons were then separated into four channels with a polarizing beam splitter and two dichroic mirrors (635DCXR, Chroma Technology). Donor photons were filtered (ET585/65m, Chroma Technology) and then focused on τ-SPAD avalanche photodiodes (PicoQuant). Acceptor photons were filtered (LP647RU and HC750/SP, Chroma Technology) and detected with SPCM-AQRH-14 single-photon avalanche diodes (PerkinElmer). The arrival times of detected photons were recorded with four channels of a HydraHarp 400 counting module (PicoQuant) with a resolution of 16 ps. All measurements were performed by exciting the donor dye with a laser power of 50 μW, measured at the back aperture of the objective. The objective was mounted on a piezo stage combination (P-733.2 and PIFOC, Physik Instrumente GmbH) to enable 3D scans. During such a scan, the sample was excited with pulsed laser light (20 MHz, 520 nm). The emitted fluorescence was used to generate fluorescence lifetime images of the scanned cells using the SymPhoTime software package (PicoQuant).

In-cell FRET and nanosecond FCS of ProTα. Single-molecule FRET efficiency histograms of ProTα in living cells were acquired by injecting samples with a protein concentration of 10 nM into the cytosol or nucleus of HeLa cells. ProTα was dissolved in the buffer used for the experiment, additionally containing 0.005% Tween 20. All experiments with ProTα were performed at 22 °C. As a sample buffer and extracellular medium, modified HBSS without Ca^{2+} and Mg^{2+} and with different NaCl concentrations was used. Besides NaCl, every buffer contained 0.3 mM Na_2HPO_4 (Sigma), 0.4 mM KH_2PO_4 (Sigma), 5.4 mM KCl (Sigma), 4.2 mM $NaHCO_3$ (Sigma) and 5.6 mM glucose (Sigma). The concentration of NaCl was adjusted to the desired final ionic strength of the buffer (30 mM, 80 mM, 137 mM, 280 mM and 980 mM NaCl), yielding buffers with total ionic strengths of 50 mM, 100 mM, 156 mM, 300 mM and 1,000 mM, respectively. The pH of the buffers was between 7.2 and 7.4. The cells were incubated in the corresponding buffer for 15 min before the sample was injected into the cells. For experiments with buffers with ionic strengths of 50 mM, 100 mM, and 156 mM, the microscopy dishes were coated with poly(L-lysine), for improved cell adhesion, before cells were added. Note that in the range of salt concentrations used here, no detectable effects on the stability and photophysics of the dyes are expected[32]. For I = 50 mM, 6 cytosolic and 1 nuclear measurements were performed; for I = 100 mM, 6 cytosolic and 8 nuclear measurements; for I = 156 mM, 8 cytosolic and 8 nuclear measurements; for I = 300 mM, 7 cytosolic and 8 nuclear measurements; for I = 1,000 mM, 5 cytosolic and 4 nuclear measurements. Each data point was collected in a different cell.

Two slightly different procedures were used for in-cell data acquisition. In both cases, before injection, a FLIM image of the cells was first taken in the axial direction for setting the z position, and then an image in the xy plane was recorded. After microinjection of sample into a few cells, success of injection was verified by recording a second FLIM image of the injected cells in the xy plane (for example, **Fig. 1b**), and the confocal volume was positioned either in the cytosol or the nucleus of the injected cells and away from regions of high cellular autofluorescence. Alternatively, in order to minimize photobleaching of the injected sample, we recorded an image of the cells only before injection, and the position of the confocal volume after injection was chosen on the basis of that image. The confocal spot was then set to the position of choice, either in the cytosol or in the nucleus of an injected cell. In both cases, positioning of the observation volume was additionally verified by the presence of fluorescence bursts (**Fig. 1c**) and the slower translational diffusion of the molecules compared to the extracellular medium (**Fig. 1d**). From each injected cell, data were acquired for 2–3 min until most of the injected protein was photobleached. The concentration of labeled protein in the injected cells was estimated by comparing the frequency of detected fluorescence bursts with reference measurements *in vitro* and yielded concentrations between 50 pM and 500 pM. (Note that at such low concentrations of injected sample, using the amplitudes of the component of FCS curves corresponding to translational diffusion for the determination of concentrations[6] is complicated by the contribution of cellular autofluorescence.)

Nanosecond FCS (nsFCS) measurements in cells were performed by injecting ProTα in HBSS without Ca^{2+} and Mg^{2+} (I = 156 mM), containing 0.005% Tween 20. Owing to the poorer photon statistics for interphoton times in the nanosecond range, the measurements require a substantially larger total number of photons to be detected than are needed for FRET efficiency histograms. Optimal signal-to-noise ratio is achieved in the range of ~1 nM of labeled protein; we thus injected correspondingly higher concentrations of labeled ProTα (20 nM). Injections and measurements were performed as described above. In contrast to the other single-molecule fluorescence measurements, nsFCS measurements were carried out using a 532-nm continuous-wave laser (LaserBoxx LMX-532S, Oxxius) at 100-μW excitation power, as measured at the back aperture of the objective. In-cell experiments were performed by measuring the fluorescence in each injected cell for 3 min. To achieve sufficient statistics, we averaged the results of 51 single-cell measurements, which resulted in a total acquisition time of 2.5 h. We note that, given the small cell-to-cell variation in $\langle E \rangle$, a large heterogeneity in dynamics seems unlikely. Also, during the measurement in one single cell, the transfer efficiency of ProTα remained constant to within experimental uncertainty (**Supplementary Fig. 13**). Control measurements in buffer were carried out with protein concentrations of 200 pM in HBSS with 0.001% Tween 20 for 12 h. Owing to the relatively low transfer efficiency of the ProTα variant used here and the limited signal-to-noise ratio, the amplitude of the cross-correlation component corresponding to chain dynamics is not detectable in our intracellular measurements (**Fig. 1f**).

In-cell FRET of Yfh1 at different temperatures. For intracellular measurements at different temperatures, a custom-made temperature-controlled cell-culture dish holder employing Peltier elements and a digital temperature controller (HAT Control-M20, BelektroniG GbR) with a Pt100 temperature sensor (**Supplementary Fig. 7**) was used. The Peltier elements were cooled by continuous water circulation. The microscope objective was temperature regulated using an aluminum collar to which a

doi:10.1038/nmeth.3475

water-cooled Peltier element was mounted[40]. For temperature-dependent control measurements in buffer, a custom-built temperature-controlled sample holder for glass cuvettes employing Peltier elements and a HAT Control-M20 temperature controller with a Pt100 temperature sensor was used as previously described[5]. For both extracellular and intracellular measurements, the temperature at the laser focus was calibrated using the temperature-dependent fluorescence lifetime of rhodamine B[40,59,60].

HeLa cells were grown in poly(L-lysine)-coated microscopy dishes. The cells were kept in HBSS without Ca^{2+} and Mg^{2+} salts (I = 156 mM). Yfh1 in cells was measured at temperatures between 279 K and 299 K. Above that temperature range, the pH became instable too quickly, which led to cell detachment (for measurements at higher temperatures, a CO_2 cabinet attached to the microscope would need to be employed to maintain the physiological pH). Nevertheless, HeLa cells were found to tolerate from ~277 K to ~303 K without loss of surface adherence after microinjection. A 50 nM Yfh1 solution in 20 mM sodium phosphate buffer, pH 7.0, and 0.005% Tween 20 was injected into HeLa cells with PLL-PEG–coated glass capillaries. Extracellular measurements were performed with 50 pM Yfh1 in 50 mM sodium phosphate buffer, pH 7.0, in HBSS and in HBSS without Ca^{2+} and Mg^{2+}. Every buffer additionally contained 144 mM β-mercaptoethanol and 0.001% Tween 20, and measurements were performed at different temperatures (for measurement in HBSS buffer, see **Supplementary Fig. 8**). The measurements at 279 K were performed in 8 individual cells, at 282 K in 11 cells, at 286 K in 8 cells, at 289 K in 10 cells, at 293 K in 10 cells, and at 299 K in 12 cells. Note that in the temperature range accessible here, only minor effects on the stability and photophysics of the dyes are expected[40,59,61]. Both here and in the cases of ProTα and GB1, some variation in the amplitudes of the donor-only peak from cell to cell was observed. Given the absence of donor-only bursts in noninjected cells (**Supplementary Fig. 14**), the most likely reasons of this variation are photobleaching and/or differences in intracellular proteolytic degradation.

Recurrence analysis of single particles (RASP) with GB1 in the cell. RASP is based on the observation that at sufficiently low sample concentrations and for timescales up to several tens of milliseconds, the probability that a fluorescently labeled molecule returns to the confocal volume is greater than the probability of a new molecule being detected[43]. Intramolecular interconversion kinetics can then be quantified by the changes in transfer efficiency histograms obtained from fluorescence bursts detected within a defined time interval that is varied systematically (**Supplementary Fig. 10**). The possibility of extracting kinetics for times longer than the diffusion time through the confocal volume (~1 ms) critically depends on the probability that two fluorescence bursts separated by a time τ were emitted by the same molecule ($p_{\mathrm{same}}(\tau)$). For the large sample volumes used in typical *in vitro* single-molecule measurements, the huge reservoir of fluorescent molecules (~10^9) leads to a rapid drop of p_{same} with increasing τ (ref. 43). However, owing to the confinement within the small cellular volume (~0.5 pl), p_{same} remains high for much longer times, as illustrated by simulations (**Fig. 4b**). Essentially, cellular confinement strongly increases the probability of observing multiple transits of a molecule through the confocal volume, which extends the effective observation time and makes recurrence analysis ideally suited for probing intracellular conformational dynamics.

RASP measurements in cells were performed at 4 °C to improve the peak separation between the unfolded and folded population (as in the case of Yfh1 (ref. 40), the unfolded state expands with decreasing temperature, resulting in lower transfer efficiency and thus better separation from the high-transfer-efficiency folded peak). For intracellular measurements, the temperature-controlled cell-culture dish holder and the objective collar were used as described above. Adherent HeLa cells for in-cell RASP experiments were grown in microscopy dishes, as described above. The medium was decanted, and the cells were washed three times with HBSS and cooled to 4 °C. Microinjection was performed using a 20 nM GB1 solution in PLL-PEG–coated glass capillaries. GB1 was injected into the cytosol of 61 HeLa cells, and fluorescence bursts were collected for 5 min per cell. To maximize throughput, we verified successful injection not by detailed fluorescence imaging but by the appearance of fluorescence bursts upon placing the confocal volume inside the cytoplasm and the reduced translational diffusion time of molecules inside cells compared to the extracellular medium. Buffer measurements were carried out in 50 mM sodium phosphate buffer, pH 7.0, 144 mM β-mercaptoethanol, and 0.001% Tween 20. Single-molecule FRET histograms in buffer were acquired in samples with a protein concentration of 20 pM. The protein concentration was chosen low enough to maximize the same-molecule probability, $p_{\mathrm{same}}(\tau)$, i.e., the probability that two fluorescence bursts, separated by a short time τ (up to 100 ms), are emitted by the same molecule that happens to return to the confocal volume. As a consequence of the low protein concentrations, fluorescence bursts had to be collected for over 20 h *in vitro* to obtain sufficient statistics for a robust analysis of the exchange dynamics between folded and unfolded states[43].

Simulation of RASP data. For a better understanding of the shape of the $p_{\mathrm{same}}(\tau)$ curve obtained from the measured data, we simulated trajectories of particles labeled with one fluorophore freely diffusing (diffusion coefficient $D = 1 \times 10^{-5}\ \mu\mathrm{m}^2/\mu\mathrm{s}$) in the confined volume of a living cell, which we approximated by a flat cylinder of diameter d = 20 µm and height h = 1, 2, 4 or 10 µm. Trajectories were simulated for each particle as a random walk with step sizes in each spatial direction drawn from a normal distribution with variance $\sigma^2 = 2D\Delta$, where Δ = 1 µs is the time between two steps. We approximate the dependence of the fluorescence photon rate n on the position (x, y, z) inside the confocal volume by a 3D Gaussian

$$n(x,y,z) = n_0 e^{-2(x^2+y^2)/\omega_1^2} e^{-2z^2/\omega_2^2}$$

with ω_1 = 0.15 µm, ω_2 = 0.9 µm, and n_0 = 0.4 photons/µs. The probability of photobleaching during one simulation step is given by $p_{\mathrm{bleach}} = \alpha_{\mathrm{bleach}} n(x, y, z)\Delta$ with $\alpha_{\mathrm{bleach}} = 0.0025$. Initially ($t = 0$), N particles were randomly distributed inside the cylinder. Each particle trajectory was simulated for 180 s or until bleaching of the fluorophore, whichever occurred first. Finally, for all times t, the combined fluorescence emission rate of all particles, $n_{\mathrm{tot}}(t)$, was calculated, and photon-detection events as a function of time were simulated with mean interphoton times given by $1/n_{\mathrm{tot}}(t)$. The simulated photon trajectories were then analyzed in the same way as the measured fluorescence data to obtain the $p_{\mathrm{same}}(\tau)$ curves presented in **Figure 4b**.

Microfluidic mixing. Microfluidic mixing experiments were performed to independently test the unfolding and folding kinetics of GB1 determined by recurrence measurements using RASP[43]. Microfluidic devices were fabricated using replica molding in poly(dimethylsiloxane) (PDMS) as described previously[62]. In concordance with the RASP measurements, microfluidic experiments were performed at 4 °C. To reach this temperature inside the microfluidic mixing device, both the microscope objective and the microfluidic device holder[62] were cooled to 1.00 °C and 0.25 °C, respectively. The actual temperature inside the microfluidic channels was confirmed by measuring the fluorescence lifetime of a 1 μM rhodamine B solution[60] (excitation power, 0.06 μW at 532 nm with 20-MHz repetition rate). To minimize surface adhesion of fluorescently labeled protein to the channel walls, we flushed microfluidic mixing devices with an aqueous solution of 0.1 mg/ml PLL(20)-g[3.5]-PEG(2) for 1 h before the actual mixing experiment. Refolding experiments were performed in the microfluidic mixing device by jumping GB1 from 800 mM to 80 mM GdmCl (**Supplementary Fig. 9**). To determine the folding-unfolding rate of GB1 starting from an ensemble biased toward the unfolded state, the center inlet of the microfluidic mixing device was filled with sample containing 250 pM GB1 in 50 mM sodium phosphate buffer with 800 mM GdmCl, 140 mM β-mercaptoethanol and 0.01% Tween 20, pH 7. Side inlets were filled with 50 mM sodium phosphate buffer, pH 7. Application of 23.4 kPa and 24.8 kPa to the center and side inlets resulted in an average flow velocity of 1 mm/s, a mixing ratio of 1:10, and a concentration of 25 pM GB1 in 80 mM GdmCl in the observation channel. Note that the pressures necessary are substantially higher than those previously reported[62] owing to increased viscosities at 4 °C, as measured using a digital viscometer (DV-I+, Brookfield).

Kinetic series were obtained by acquiring transfer efficiency histograms at different positions along the observation channel, corresponding to different times after mixing. 50 μW of excitation power of a 532-nm continuous-wave laser (LaserBoxx LMX-532S, Oxxius) were used. Transfer efficiency histograms were recorded for 5 min per data point. Three sets of measurements were averaged, and for the estimation of error bars, we assumed the uncertainties to be uniformly distributed. Positions along the observation channel were converted into times after mixing as described previously[63]. Positions corresponding to short times after mixing (up to 50 ms) were accessed using a piezo stage combination (P-733.2 and PIFOC, PI) and those for longer times were accessed using the manual translation stage of the microscope. Histograms were obtained from 2.1 ms to 377 ms after mixing.

Data analysis. *Transfer efficiency histograms.* Fluorescence bursts from individual molecules were identified by combining successive photons separated by interphoton times of less than 100 μs and retained as a burst if the total number of photons detected was larger than a threshold of 30–50, chosen depending on the intensity of the background signal. The background was variable during a measurement owing to photobleaching and was quantified by splitting the time traces into short intervals where the background could be assumed to be constant. Detected bursts from each interval were summed up to construct the final transfer efficiency histogram. The transfer efficiency of each burst was calculated according to $E = n_A/(n_A + n_D)$, where n_D and n_A are the numbers of donor and acceptor photons, respectively, corrected for background in the corresponding interval, acceptor direct excitation, channel cross-talk, differences in detector efficiencies, and quantum yields of the dyes as previously described[64]. (Note that owing to these corrections, E values >1 and <0 are possible in the transfer efficiency histograms.)

For all three proteins investigated here, the 'donor-only' population (i.e., the fraction of fluorescence bursts from molecules lacking an active acceptor dye with an apparent transfer efficiency close to 0) is larger for in-cell measurements than for *in vitro* measurements. Several indications suggest that this signal originates from a preferential inactivation of the acceptor dye in the cell or from cellular degradation of part of the injected molecules, as follows. (i) The yield of correctly FRET-labeled molecules for all samples was >90% (see "Protein expression and labeling"), which is reflected by the small donor-only population in *in vitro* measurements (**Figs. 1e** and **4c** and **Supplementary Fig. 8**). (ii) Measurements in noninjected cells analyzed with the thresholds of our in-cell measurements yielded a negligible signal in the transfer efficiency range of the donor-only peak (**Supplementary Fig. 14**). (iii) The fluorescence lifetime of the donor-only population was not significantly different from the donor lifetime in extracellular measurements (**Supplementary Fig. 5**), suggesting that the donor-only signal does not originate from cellular autofluorescence, which exhibits a different lifetime (**Fig. 1b** and **Supplementary Fig. 3**). (iv) The lower number of fluorescence bursts with a transfer efficiency close to 0 upon scanning the sample compared to stationary measurements (**Supplementary Fig. 4**) indicates the contribution of photobleaching to the donor-only population.

Fitting of transfer efficiency histograms. To analyze subpopulations in transfer efficiency histograms, we approximated peaks in the histograms with Gaussian (G) and four-parameter log-normal (L) peak functions for symmetric and asymmetric peaks, respectively:

$$G(E) = Ae^{-\frac{(E-E_0)^2}{2w^2}}$$

and

$$L(E) = A\exp\left(-\frac{\ln(2)}{\ln(a)^2}\ln\left(1+\frac{a^2-1}{aw}(E-E_0)\right)^2\right)$$

with E being the transfer efficiency, A the peak amplitude, E_0 the peak position, w the peak width, and a the asymmetry of the peak[65–67]. For fitting more than one peak, the histogram was analyzed with a sum of peak functions. In many cases where several histograms contained peaks from the same species, the peaks were described with a global fit over all histograms, with E_0, w, and/or a for some subpopulations as shared or fixed parameters (amplitudes were never constrained). Where necessary, the choice of fixed parameter values was based on histograms where the respective peaks were as well defined as possible (i.e., separated from other subpopulations and/or with the best available statistics, or from recurrence histograms in the case of GB1).

doi:10.1038/nmeth.3475

Specifically, transfer efficiency histograms of ProTα (**Figs. 1** and **2** and **Supplementary Fig. 13**) were fitted using a log-normal peak function for the donor-only population with a fixed width (0.15) and asymmetry (1.32). These values were chosen to best describe the observed shape of the donor-only population. The FRET-labeled population was fitted with a Gaussian peak function without constraints of fit parameters.

Transfer efficiency histograms of frataxin (**Fig. 3** and **Supplementary Fig. 8**) were fitted globally using a log-normal peak function for the donor-only population, with a fixed width (0.14 *in vitro*, 0.2 *in vivo*), and asymmetry (1.1 *in vitro*, 1.4 *in vivo*); the position was a fit parameter shared by all histograms of the respective data set. These values were chosen to best describe the observed shape of the donor-only population. The unfolded population was fitted with a Gaussian peak function using the width as a fit parameter shared by all histograms in the case of **Supplementary Figure 8** and, owing to the lower statistics, with a width fixed to 0.09 in case of **Figure 3**. The folded population was fitted with a Gaussian peak function using the width as a global fit parameter for all histograms in the case of **Supplementary Figure 8** and with fixed width (0.09) in case of **Figure 3**. In the case of **Figure 3**, the position of the native population was a fit parameter shared by all histograms (based on the assumption that the transfer efficiency of the folded population is virtually constant over this temperature range, as *in vitro*). The widths of the unfolded and folded populations *in vivo* (0.09) were chosen on the basis of observations made with ProTα and GB1, for which a slight broadening of peaks was observed compared to *in vitro*.

Transfer efficiency histograms of GB1 (**Fig. 4** and **Supplementary Figs. 9** and **10**) were fitted globally using a log-normal peak function for the donor-only and folded populations and using a Gaussian peak function for the unfolded population. In all cases, the donor-only population was fitted with fixed position (−0.03), width (0.14), and asymmetry (1.09). All parameters for the donor-only, folded, and unfolded subpopulations (and for ProTα in the control) were obtained from globally fitting the early-time recurrence histograms of each population without any constraints but using p, w, and a of the respective populations as shared parameters; the results from these fits were used to fix the parameters in the remaining recurrence histograms. Correspondingly, the unfolded population was fitted with a fixed position (0.85) and width (0.08 *in vitro*, 0.1 *in vivo*) (based on ΔE = (0.7–0.9) and ΔT = (0–1 ms)). The native population was fitted with a fixed position (1.0), width (0.07 *in vitro*, 0.09 *in vivo*), and asymmetry (0.42 *in vitro*, 0.63 *in vivo*) (based on ΔE = (0.95–1.1) and ΔT = (0–1 ms)). The histograms shown in **Supplementary Figure 11** were fitted with fixed position (−0.05), width (0.13), and asymmetry (1.17) for the donor-only population (log-normal peak). The population of ProTα (Gaussian peak) was fitted with fixed position (0.43) and width (0.1) (based on ΔE = (0.35–0.55) and ΔT = (0–5 ms)). The population of GB1 wild type was fitted with fixed position (1.0), width (0.09), and asymmetry (0.76), using log-normal peak functions (based on ΔE = (0.9–1.1) and ΔT = (0–5 ms)).

Interdye distances and radii of gyration. Dimensions of unfolded/intrinsically disordered proteins were calculated from the measured transfer efficiencies as described earlier[66]. Briefly, the mean transfer efficiency, $\langle E\rangle$, can be expressed as a function of the distance dependence of the transfer efficiency, $E(r)$, weighted for the dye-to-dye distance distribution, $P(r)$, sampled by the chain,

$$\langle E\rangle = \int_a^{l_c} E(r)P(r)\mathrm{d}r$$

with

$$E(r) = \frac{1}{1+(r/R_0)^6}$$

where a is the distance of closest approach of the two dyes, l_c is the contour length of the labeled protein segment, and R_0 is the Förster radius of the chosen dye pair. To enable a direct comparison with previous experiments *in vitro*[32,40,66], we assumed a simple Gaussian chain model for the distribution of distances, $P(r)$,

$$P(r) = 4\pi r^2\left(\frac{3}{2\pi\langle r^2\rangle}\right)^{3/2}\exp\left(-\frac{3r^2}{2\langle r^2\rangle}\right)$$

where $\langle r^2\rangle$ is the mean squared interdye distance. These equations were solved for $\langle r^2\rangle$ numerically and then converted to the mean radius of gyration, R_g, of the chain using

$$R_g^2 = \frac{1}{6}\langle r^2\rangle.$$

We note that a systematic uncertainty in the Förster radius propagates linearly to $\langle r^2\rangle^{1/2}$ and thus introduces a systematic uncertainty in the absolute values of R_g. However, as we only compare measurements between identical samples under different conditions, this contribution is not considered here.

Fluorescence anisotropy values of donor emission of all samples were in the range of 0.04–0.11 *in vitro* and 0.10–0.17 in in-cell experiments, as quantified on the basis of the polarization selectivity of our single-molecule measurements (**Supplementary Fig. 6**), supporting the approximation $\kappa^2 = 2/3$ used for the calculation of the Förster radii. Donor fluorescence lifetimes of ProTα in the absence of energy transfer resulted in values of 3.8 ns in buffer, 3.8 ns in the cytosol, and 3.9 ns in the nucleus (**Supplementary Fig. 5**), indicating only minor effects of the respective environment on the quantum yield and photophysics of the donor dye.

Fluorescence correlation spectroscopy. To extract the diffusion time of the labeled proteins, we correlated[25] the fluorescence intensity fluctuations measured in FRET experiments according to

$$g_{ij}(\tau) = \frac{\langle \delta n_i(0)\delta n_j(\tau)\rangle}{\langle n_i\rangle^2}$$

with i, j = A, D and where $n_i(0)$ and $n_j(\tau)$ are the fluorescence count rates for channels i and j at time 0 and after a lag time τ, respectively, and $\delta n_{i,j} = n_{i,j} - \langle n_{i,j}\rangle$ are the corresponding deviations from the mean count rates. The duration of the single-cell FRET measurements allows correlations to be calculated to a minimum lag time of 10^{-5} s, with sufficient photon statistics to quantify the photophysics of the triplet state and the translational diffusion time of the molecule through the confocal volume. The resulting cross-correlation curves were fitted with[68]

$$g_{\mathrm{AD}}(\tau)=1+\frac{1}{N_{\mathrm{AD}}}\frac{1+c_{\mathrm{T}}^{(\mathrm{AD})}e^{-|\tau|/\tau_{\mathrm{T}}^{(\mathrm{AD})}}}{(1+|\tau|/\tau_{\mathrm{D}})(1+|\tau|/s^{2}\tau_{\mathrm{D}})^{1/2}}$$

where the term in the numerator describes triplet blinking with amplitude c_{T} and a decay time τ_{T}. The denominator describes the diffusion of a protein through the confocal volume, with N_{AD} denoting the mean number of doubly labeled protein molecules in the confocal volume, s the ratio of axial over lateral radius of the confocal volume, and τ_{D} the translational diffusion time.

nsFCS. Donor and acceptor autocorrelation curves and cross-correlation curves between acceptor and donor channels were calculated as described earlier[25,69]. The resulting auto- and cross-correlation curves were fitted globally up to a delay time of 1 µs using

$$g_{ij}(\tau)=1+\frac{1}{N}\left(1-c_{\mathrm{AB}}e^{-|t-t_{0}|/\tau_{\mathrm{AB}}}\right)\left(1+c_{\mathrm{CD}}e^{-|t-t_{0}|/\tau_{\mathrm{CD}}}\right)\left(1+c_{\mathrm{T}}e^{-|t-t_{0}|/\tau_{\mathrm{T}}}\right)$$

with i, j = A, D and with N denoting the mean number of proteins in the confocal volume. The three multiplicative terms describe photon antibunching (AB), chain dynamics (CD), and triplet blinking (T) of the dyes. τ_{CD} was extracted from the global fit of all three correlations (with τ_{CD} as a shared fit parameter) to quantify chain reconfiguration dynamics of ProTα.

RASP analysis. The fluorescence signal from the first 50 s of data acquisition in each cell showed a significant drop in signal in both donor and acceptor channels because of photobleaching and was excluded from the analysis. The fluorescence bursts identified in the remaining 250 s for 61 individual HeLa cells were combined and then subjected to RASP analysis. RASP analysis and the calculation of the p_{same} function from the burst time correlation function was carried out as described previously[43]. p_{same} can, under equilibrium conditions, be calculated from the time correlation of burst events, $g_{\mathrm{bursts}}(\tau)$, by the equation $p_{\mathrm{same}}(\tau) = 1 - 1/g_{\mathrm{bursts}}(\tau)$. Inside the cells, however, we observe a decrease of fluorescent particle concentration over time caused by photobleaching, which means that for a given τ, $p_{\mathrm{same}}(\tau)$ is increasing over time. Hence, the $p_{\mathrm{same}}(\tau)$ calculated by the equation above from all observed bursts is only an average value. Recurrence FRET efficiency histograms were obtained by first selecting photon bursts from a small transfer efficiency range ($\Delta E = (0.7–0.9)$ for selecting mainly unfolded GB1; $\Delta E = (0.95 - 1.1)$ for selecting mainly folded GB1) and then building the FRET efficiency histogram only from bursts detected within a short time (the recurrence interval, T; **Supplementary Fig. 10**) after those initial bursts. Systematic variation of the recurrence interval allows the kinetics of interconversion between the subpopulations to be determined essentially in a model-free manner[43]. Note that the choice of the initial transfer efficiency range does not affect the relaxation time but only the amplitude of the transient observed[43] (**Supplementary Fig. 15**). The fractions of folded molecules from recurrence FRET efficiency histograms starting from unfolded (ΔE_1) and folded molecules (ΔE_2) using recurrence analysis over a 50-ms time window were fitted with a model including the folding-unfolding dynamics of recurring molecules, which allows the relaxation time of this process to be extracted[43]. The equilibrium constant was obtained from the ratio of unfolded and folded fractions from the total transfer efficiency histogram of a RASP measurement (**Fig. 4**).

Microfluidic mixing. Detected photons were binned into windows of 1 ms, and corrections for background, differences in quantum yields of the dyes, detection efficiencies, and cross-talk were applied. Bursts were identified using a threshold for the total number of detected photons per bin. Given the pronounced decrease in flow velocity in the first part of the observation channel corresponding to the entrance length, the threshold was varied as a function of the flow velocity from 21 photons per bin for the first recorded data point to 35 photons per bin for longer times after mixing to account for the differences in residence time. Transfer efficiency histograms show three peaks, corresponding to molecules lacking an active acceptor dye, unfolded GB1 ($E = 0.86$) and folded GB1 ($E = 0.99$), respectively. Histograms were fitted globally with a Gaussian and a log-normal peak function for the unfolded and folded populations, respectively; peak positions and widths were set as shared parameters and amplitudes fitted individually. The fraction of folded molecules was calculated from the relative area of the folded peak as obtained from the fits. The relaxation time of folding was extracted from a single exponential fit to the mean of three measurements. The equilibrium constant, K_{eq}, for the final conditions was obtained from a transfer efficiency histogram of GB1 measured in a cuvette in 50 mM sodium phosphate buffer with 80 mM GdmCl, 140 mM β-mercaptoethanol, and 0.01% Tween 20, pH 7. K_{eq} and the relaxation time from the mixing experiment were extrapolated to 0 M GdmCl using an m-value for folding of 1.5 kcal/mol/M and an m-value for unfolding of 0.4 kcal/mol/M. The m-values were average values taken from m-values of the single mutants F30L and G41A and wild-type GB1 (ref. 56).

Code availability. Data analysis was performed with a custom module for Mathematica (Wolfram Research), which is available upon request.

48. van de Linde, S., Heilemann, M. & Sauer, M. Live-cell super-resolution imaging with synthetic fluorophores. *Annu. Rev. Phys. Chem.* **63**, 519–540 (2012).
49. Raj, A. & van Oudenaarden, A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216–226 (2008).
50. Costantini, L.M., Fossati, M., Francolini, M. & Snapp, E.L. Assessing the tendency of fluorescent proteins to oligomerize under physiologic conditions. *Traffic* **13**, 643–649 (2012).
51. Plochowietz, A., Crawford, R. & Kapanidis, A.N. Characterization of organic fluorophores for *in vivo* FRET studies based on electroporated molecules. *Phys. Chem. Chem. Phys.* **16**, 12688–12694 (2014).
52. Kolmakov, K. *et al.* Red-emitting rhodamines with hydroxylated, sulfonated, and phosphorylated dye residues and their use in fluorescence nanoscopy. *Chemistry* **18**, 12986–12998 (2012).
53. Zanetti-Domingues, L.C., Tynan, C.J., Rolfe, D.J., Clarke, D.T. & Martin-Fernandez, M. Hydrophobic fluorescent probes introduce artifacts into single molecule tracking experiments due to non-specific binding. *PLoS ONE* **8**, e74200 (2013).
54. Yi, S., Brickenden, A. & Choy, W.Y. A new protocol for high-yield purification of recombinant human prothymosin alpha expressed in *Escherichia coli* for NMR studies. *Protein Expr. Purif.* **57**, 1–8 (2008).
55. Gronenborn, A.M. *et al.* A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science* **253**, 657–661 (1991).

doi:10.1038/nmeth.3475

56. McCallister, E.L., Alm, E. & Baker, D. Critical role of β-hairpin formation in protein G folding. *Nat. Struct. Biol.* **7**, 669–673 (2000).
57. Kenausis, G.L. *et al.* Poly(L-lysine)-*g*-poly(ethylene glycol) layers on metal oxide surfaces: attachment mechanism and effects of polymer architecture on resistance to protein adsorption. *J. Phys. Chem. B* **104**, 3298–3309 (2000).
58. Nicoletti, I., Migliorati, G., Pagliacci, M.C., Grignani, F. & Riccardi, C. A rapid and simple method for measuring thymocyte apoptosis by propidium iodide staining and flow cytometry. *J. Immunol. Methods* **139**, 271–279 (1991).
59. Nettels, D. *et al.* Single-molecule spectroscopy of the temperature-induced collapse of unfolded proteins. *Proc. Natl. Acad. Sci. USA* **106**, 20740–20745 (2009).
60. Benninger, R.K. *et al.* Quantitative 3D mapping of fluidic temperatures within microchannel networks using fluorescence lifetime imaging. *Anal. Chem.* **78**, 2272–2278 (2006).
61. Wuttke, R. *et al.* Temperature-dependent solvation modulates the dimensions of disordered proteins. *Proc. Natl. Acad. Sci. USA* **111**, 5213–5218 (2014).
62. Wunderlich, B. *et al.* Microfluidic mixer designed for performing single-molecule kinetics with confocal detection on timescales from milliseconds to minutes. *Nat. Protoc.* **8**, 1459–1474 (2013).
63. Wunderlich, B., Nettels, D. & Schuler, B. Taylor dispersion and the position-to-time conversion in microfluidic mixing devices. *Lab Chip* **14**, 219–228 (2014).
64. Schuler, B., Müller-Späth, S., Soranno, A. & Nettels, D. Application of confocal single-molecule FRET to intrinsically disordered proteins. *Methods Mol. Biol.* **896**, 21–45 (2012).
65. Schuler, B., Lipman, E.A. & Eaton, W.A. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature* **419**, 743–747 (2002).
66. Hoffmann, A. *et al.* Mapping protein collapse with single-molecule fluorescence and kinetic synchrotron radiation circular dichroism spectroscopy. *Proc. Natl. Acad. Sci. USA* **104**, 105–110 (2007).
67. Kuzmenkina, E.V., Heyes, C.D. & Nienhaus, G.U. Single-molecule Forster resonance energy transfer study of protein dynamics under denaturing conditions. *Proc. Natl. Acad. Sci. USA* **102**, 15471–15476 (2005).
68. Zander, C., Enderlein, J. & Keller, R.A. *Single Molecule Detection in Solution: Methods and Applications* (Wiley, 2003).
69. Nettels, D., Hoffmann, A. & Schuler, B. Unfolded protein and peptide dynamics investigated with single-molecule FRET and correlation spectroscopy from picoseconds to seconds. *J. Phys. Chem. B* **112**, 6137–6146 (2008).

doi:10.1038/nmeth.3475

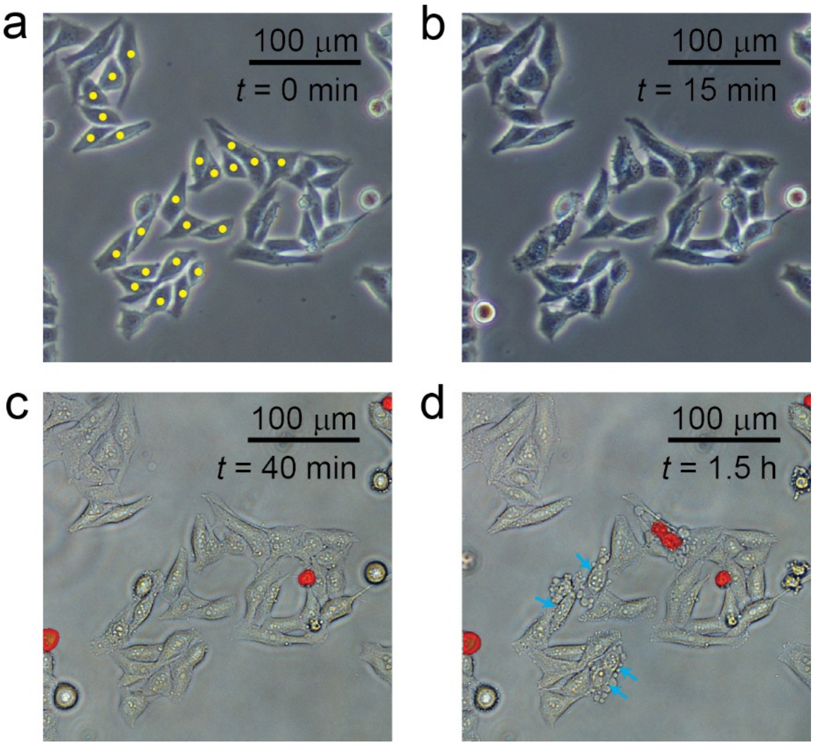
Supplementary Information

Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells

Iwo König[1], Arash Zarrine-Afsar[1], Mikayel Aznauryan[1], Andrea Soranno, Bengt Wunderlich, Fabian Dingfelder, Jakob C Stüber, Andreas Plückthun, Daniel Nettels, and Benjamin Schuler

[1] These authors contributed equally to this work

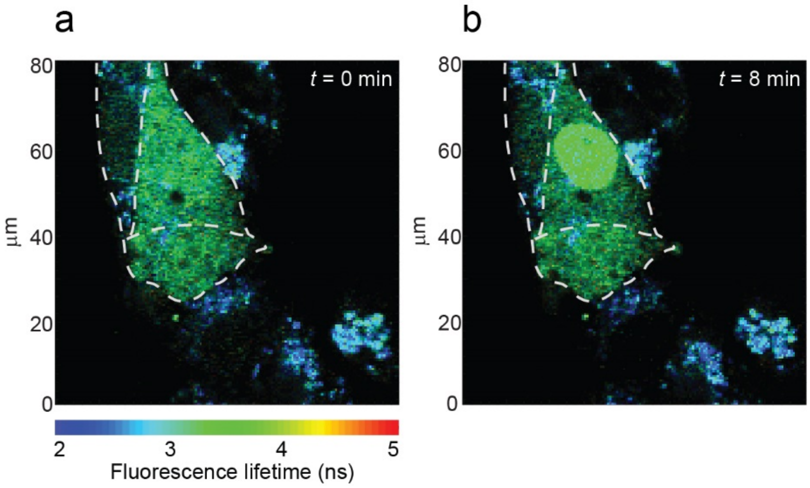
(Nature Methods, 2015)



Supplementary Figure 1

Testing cell integrity after microinjection.

(**a**) HeLa cells before injection at pH 7.2, in HBSS buffer. Cells that were to be injected are indicated by a yellow dot. (**b**) The same cells 15 min after microinjection of 20 mM sodium phosphate buffer (0.5 bar, 50 ms pulse duration, one injection per cell); only minor morphological changes of the injected cells are visible. (**c**) 40 min after injection; an overlay of a phase contrast image and the fluorescence image after addition of propidium iodide to the culture medium shows no signs of cell blebbing or cell death of the injected cells. The three red fluorescent spots originate from dead cells already present before injection. (**d**) 1.5 h after injection; four cells (18% of injected cells) formed blebs (indicated by blue arrows), and one dead cell is visible upon propidium iodide staining.



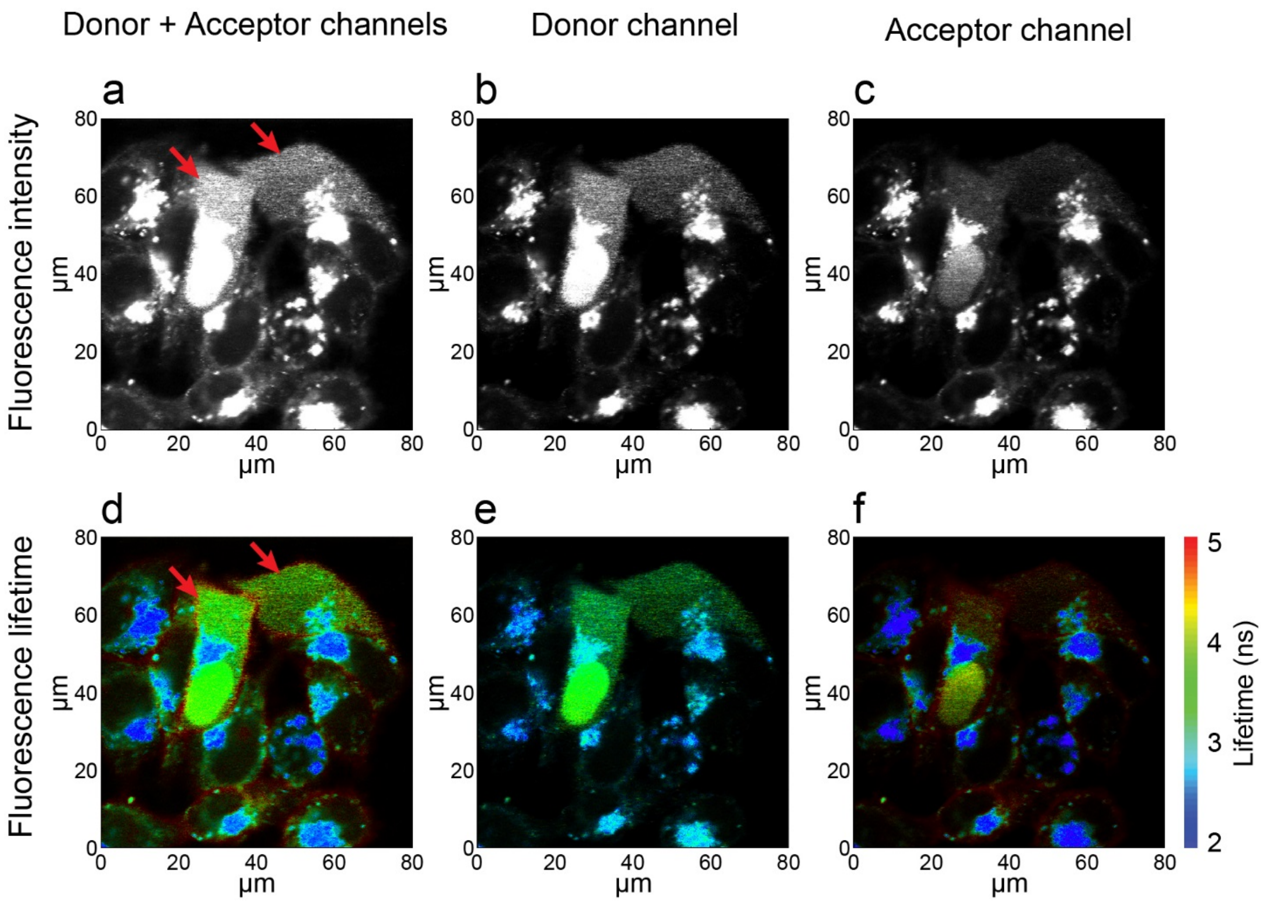
Supplementary Figure 2

Nuclear localization of ProTα observed after microinjection.

(**a**) FLIM image showing the uniform distribution of ProTα directly after injection of ProTα into the cytosol of three different HeLa cells (cell boundaries indicated by dashed lines). Due to differences in injection efficiency, the ProTα concentration in the left cell is slightly lower than in the other two cells. (**b**) 8 min later, an accumulation of ProTα in the nucleus of the middle cell is visible, indicating active nuclear transport[1]. The absence of nuclear accumulation in some cells may be due to cell cycle differences[2].

References:

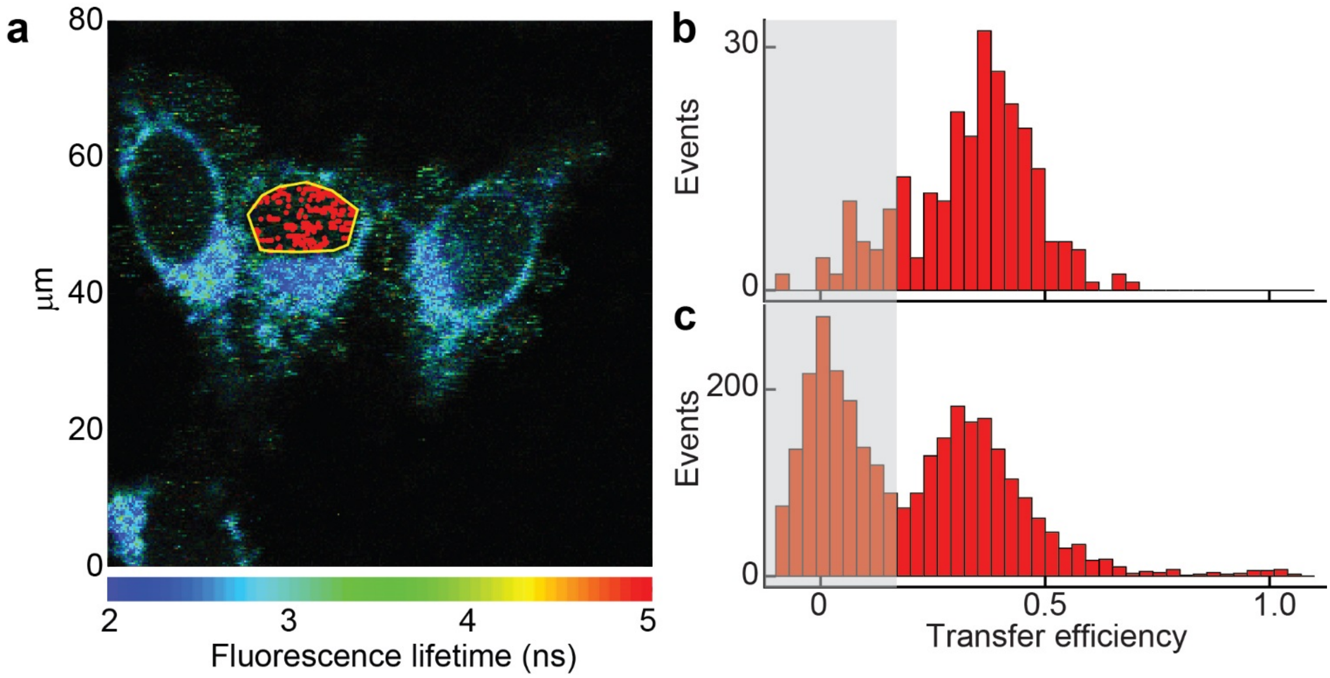
1. Manrow, R.E., Sburlati, A.R., Hanover, J.A. & Berger, S.L. Nuclear targeting of prothymosin alpha. *J. Biol. Chem.* **266**, 3916-3924 (1991).
2. Martin-Aparicio, E., Avila, J. & Lucas, J.J. Nuclear localization of N-terminal mutant huntingtin is cell cycle dependent. *Eur. J. Neurosci.* **16**, 355-359 (2002).



Supplementary Figure 3

Comparison of fluorescence intensity and fluorescence lifetime images.

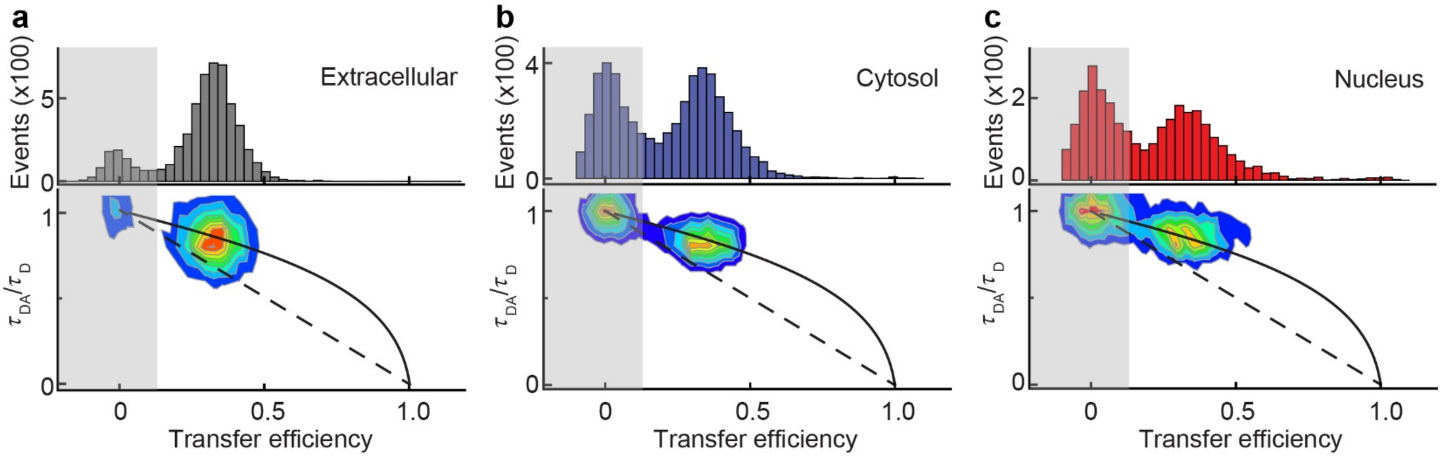
This array of images compares fluorescence intensity and lifetime images of HeLa cells. The same scan is shown in **Figure 1b**, where two cells were injected with fluorescently labeled ProTα (red arrows in (**a**) and (**d**)). The upper row shows only fluorescence intensities, with (**a**) showing donor and acceptor channels combined, (**b**) only the donor channel, and (**c**) only the acceptor channel. The lower row additionally uses the fluorescence lifetime information, with (**d**) donor and acceptor channels combined, (**e**) only the donor channel, and (**f**) only the acceptor channel (color code to the right). The comparison illustrates that fluorescence intensity information is sufficient to distinguish injected from non-injected cells, but lifetime information can be helpful additionally, e.g. to aid the distinction of fluorescence from injected sample and cellular autofluorescence (if differences in lifetimes exist).



Supplementary Figure 4

Extracting FRET efficiency histograms of intracellular ProTα from area scans.

To clarify whether an immobile subpopulation with different conformational properties from freely diffusing molecules was present, the confocal volume was scanned over an area of 80 by 80 µm in the cell during data acquisition. (**a**) Area scan showing HeLa cells after injection of FRET-labeled ProTα in the cytosol. The image was recorded with a resolution of 256 x 256 pixels and a scanning speed of 0.5 pixel/ms. The color of each pixel indicates the fluorescence lifetime. The yellow polygon indicates the area (nucleus of the middle cell) where burst detection was carried out. The red dots within the polygon indicate the positions of photon bursts detected during scanning, which were used to construct the FRET efficiency histogram in (**b**). (**b**) FRET efficiency histogram of ProTα constructed from the FLIM image. The average transfer efficiency determined from area scans in the nuclei of three different cells was 0.34 ± 0.03. The population at zero transfer efficiency (shaded) originates from molecules lacking active acceptor chromophores. (**c**) The sum of all transfer efficiency histograms of ProTα measured in the nucleus of HeLa cells without scanning shows a transfer efficiency of 0.35 ± 0.03, indicating the absence of immobile ProTα with a transfer efficiency different from freely diffusing molecules. Note also that the lower number of events with a transfer efficiency close to zero in (**b**) compared to (**c**) indicates a pronounced contribution of photobleaching to the donor-only population.



Supplementary Figure 5

Fluorescence lifetime vs. transfer efficiency histograms of ProTα.

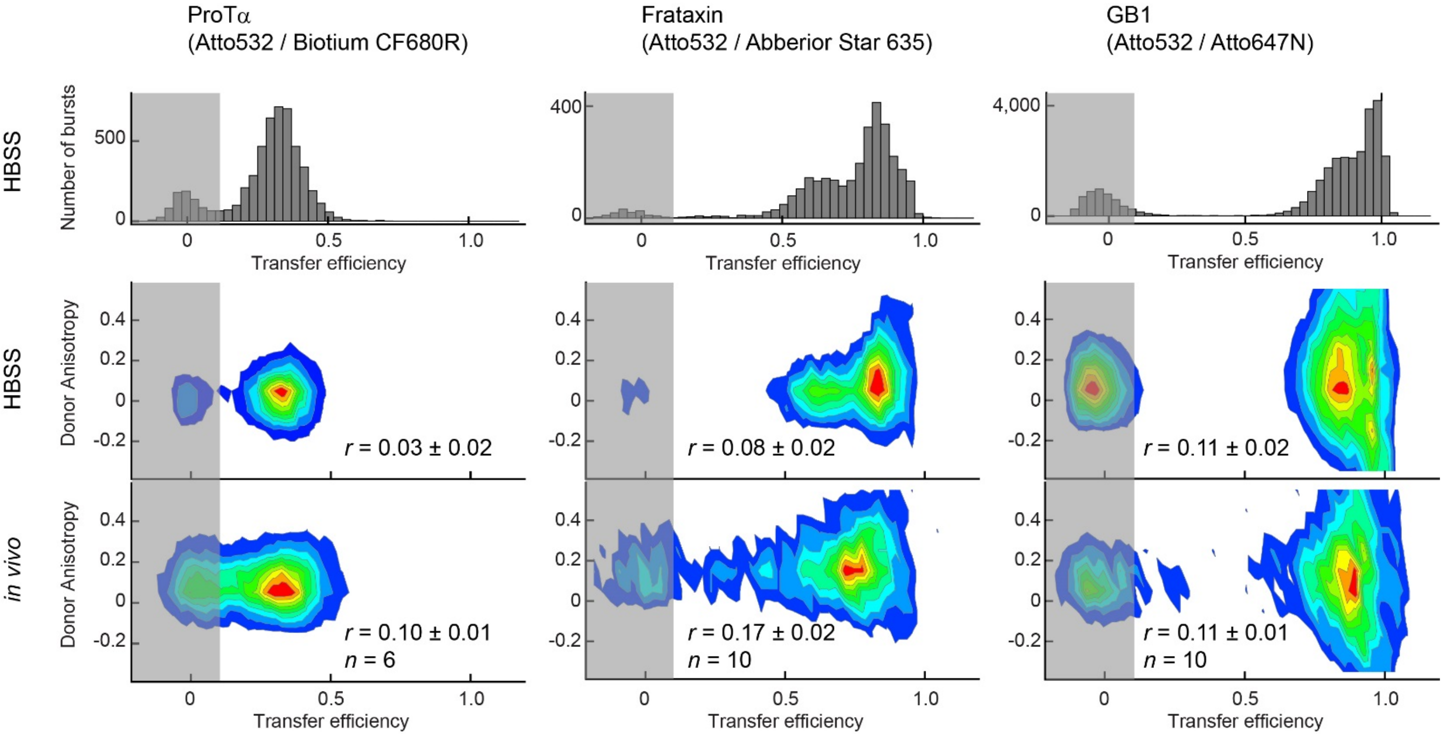
Transfer efficiency histograms (top) and the corresponding two-dimensional histograms of relative donor fluorescence lifetime (τ_{DA}/τ_D, where τ_{DA} is the lifetime in the presence and τ_D in the absence of acceptor) versus transfer efficiency (bottom) of ProTα (**a**) in the extracellular medium; (**b**) the sum of all histograms measured in the cytosol (n = 8); and (**c**) in the nucleus (n = 8) of HeLa cells incubated in a cell culture medium with *I* = 156 mM. The dashed line shows the expected relation between the average values of τ_{DA}/τ_D and transfer efficiency for a fixed distance between the two fluorophores, the solid line the relation for a chain sampling a broad distribution of distances corresponding to a Gaussian chain[3] (Online Methods). The population of unfolded ProTα is close to the solid line in all three cases, indicating that ProTα is rapidly reconfiguring not only in buffer[3] but also within the cells. The peaks at zero transfer efficiency (shaded) are due to molecules lacking an active acceptor dye and serve as a reference for determining τ_D.

Note that the peak at zero transfer efficiency is larger for intracellular (**b**,**c**) compared with the extracellular measurement (**a**), but the contribution of intracellular background to the “donor-only” peak is small, as indicated by measurements in non-injected cells (**Supplementary Fig. 14**) and in agreement with the high yields of doubly-labeled protein (Online Methods). Additionally, the absolute fluorescence lifetime of this “donor-only” population is not significantly different from the donor lifetime in extracellular measurements (3.8 ns in extracellular medium, 3.8 ns in the cytosol, and 3.9 ns in the nucleus) and is thus most likely to be due to a preferential inactivation of the acceptor dye by photobleaching (see also **Supplementary Fig. 4**) or to cellular degradation of part of the injected molecules.

We note that it will be interesting to investigate variations in biomolecular conformation or folding mechanism by positioning the observation volume in different regions within the cell or on the plasma membrane (within the spatial resolution of the diffraction-limited confocal volume), e.g. for DNA-binding proteins that fold upon binding their target in the nucleus, or for membrane proteins, respectively.

References:

3. Soranno, A. et al. Quantifying internal friction in unfolded and intrinsically disordered proteins with single molecule spectroscopy. *Proc. Natl. Acad. Sci. USA* **109**, 17800-17806 (2012).



Supplementary Figure 6

Comparison of donor fluorescence anisotropies from confocal single-molecule measurements *in vitro* and *in vivo*.

Donor fluorescence anisotropies for ProTα (labeled with Atto 532/Biotium CF680R), frataxin (Atto 532/Abberior Star 635), and GB1 (Atto 532/Atto 647N) upon excitation at 520 nm from confocal single-molecule measurements *in vitro* (in HBSS, histograms and first row of donor anisotropy vs. transfer efficiency plots) and in HeLa cells (*in vivo*, bottom row). The areas shaded in gray indicate the donor-only population and were not used for the analysis of the donor anisotropy. The donor anisotropies, *r*, were calculated for every fluorescence burst with a transfer efficiency greater than 0.2 using[4-6]

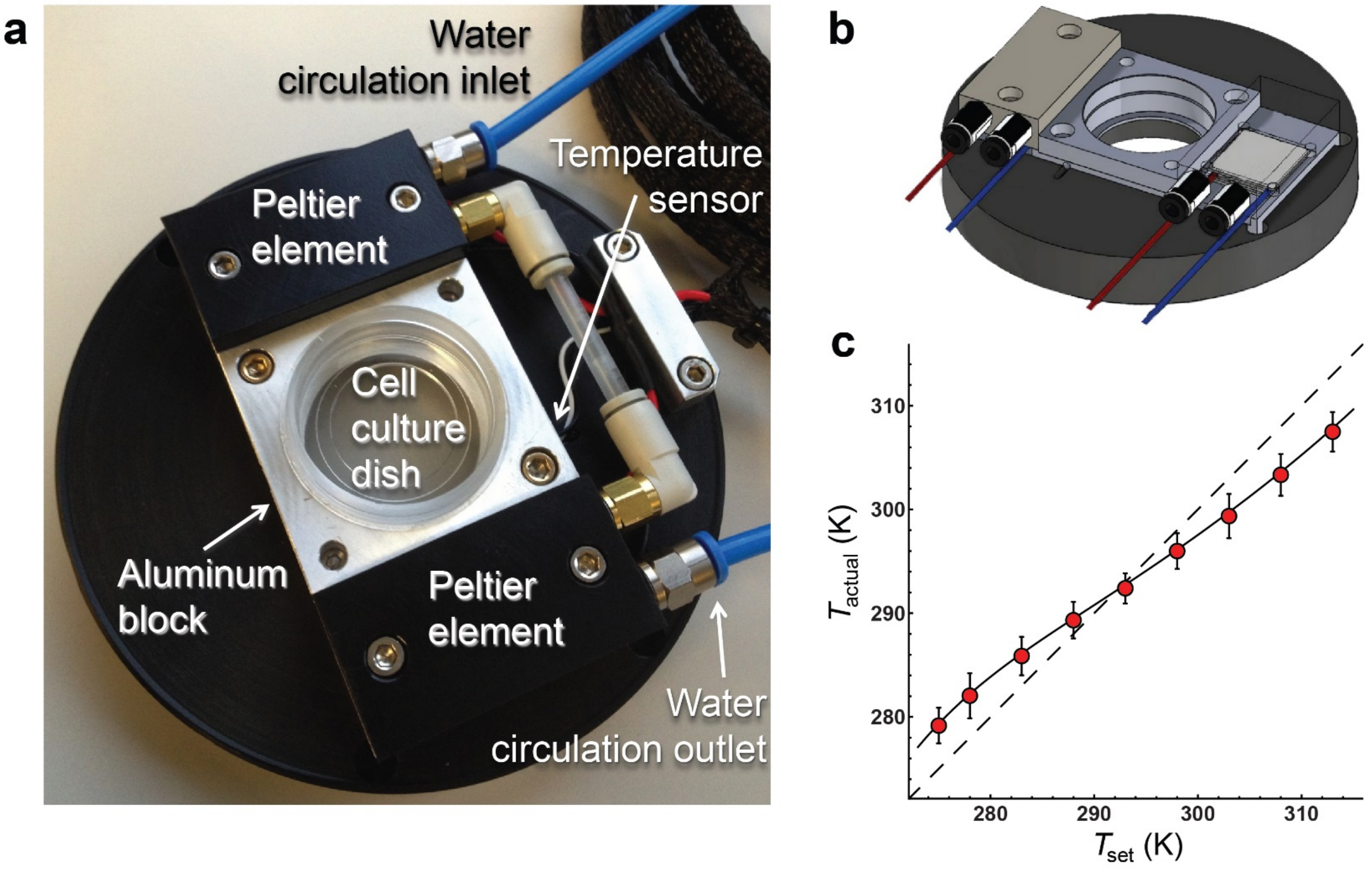
$$r = \frac{n_p - G \cdot n_s}{(1-3L_2)n_p + (2-3L_1)G \cdot n_s},$$

where $n_{p,s}$ are the numbers of donor photons in a burst with parallel (*p*) and perpendicular polarization (*s*) with respect to the polarization of the excitation light; *G* is a correction factor for differences in detection efficiencies in the parallel and the perpendicular donor channels ($G_{Atto532/CF680R}$ = 1.08, $G_{Atto532/AS635}$ = 1.11, $G_{Atto532/Atto647N}$ = 1.11), and $L_{1,2}$ are factors correcting primarily for the effect of the high numerical aperture lens used for excitation and detection[4, 5] (L_1 = 0.098, L_2 = 0.094)[6]. Mean anisotropies were calculated by combining all photons from bursts with $E > 0.2$ and calculating *r* as shown above.

Donor fluorescence anisotropies in buffer were found to be 0.03 ± 0.02 for ProTα, 0.08 ± 0.02 for frataxin, and 0.11 ± 0.02 for GB1. The uncertainties in the values in buffer were estimated by error propagation assuming an uncertainty in *G* of 5%. Anisotropies *in vivo* were higher than in buffer for ProTα (0.10 ± 0.01, mean and standard deviation from 6 cells) and frataxin (0.17 ± 0.02, mean and standard deviation from 10 cells). In the case of GB1, the anisotropy *in vivo* (0.11 ± 0.01, mean and standard deviation from 10 cells) was not significantly higher than *in vitro*.

References:

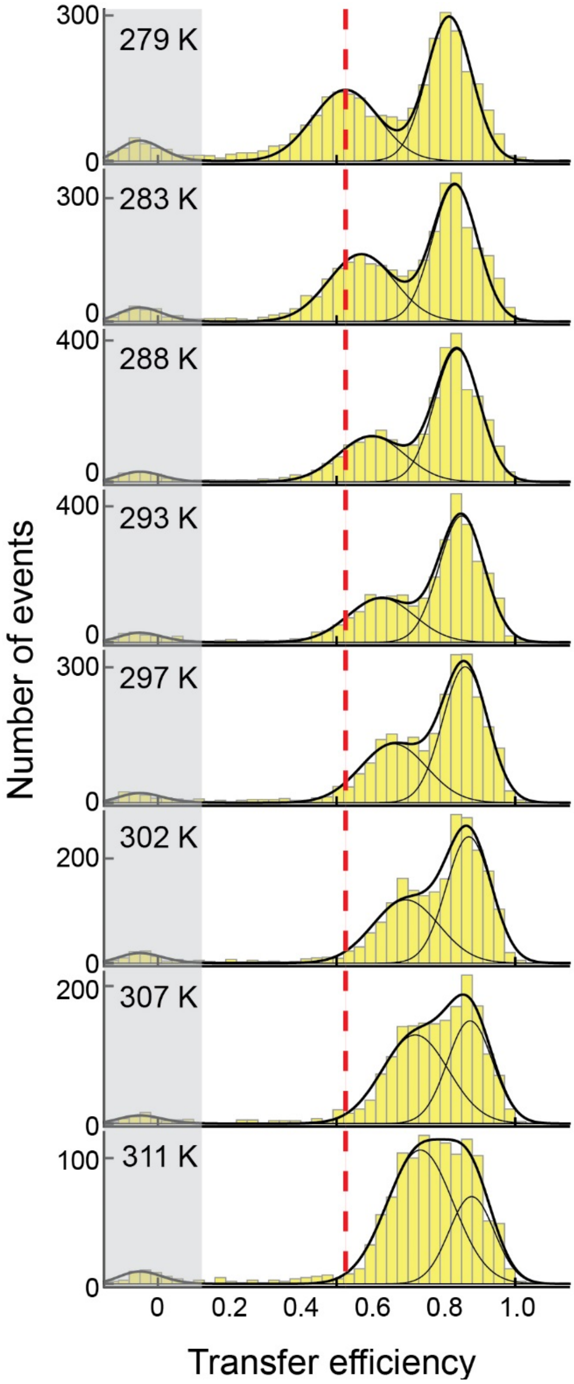
4. Koshioka, M., Sasaki, K. & Masuhara, H. Time-Dependent Fluorescence Depolarization Analysis in 3-Dimensional Microspectroscopy. *Appl. Spectrosc.* **49**, 224-228 (1995).
5. Schaffer, J. et al. Identification of single molecules in aqueous solution by time-resolved fluorescence anisotropy. *J Phys Chem A* **103**, 331-336 (1999).
6. Kellner, R. et al. Single-molecule spectroscopy reveals chaperone-mediated expansion of substrate protein. *Proc. Natl. Acad. Sci. USA* **111**, 13355-13360 (2014).



Supplementary Figure 7

Custom-built temperature-controlled holder for cell culture dishes.

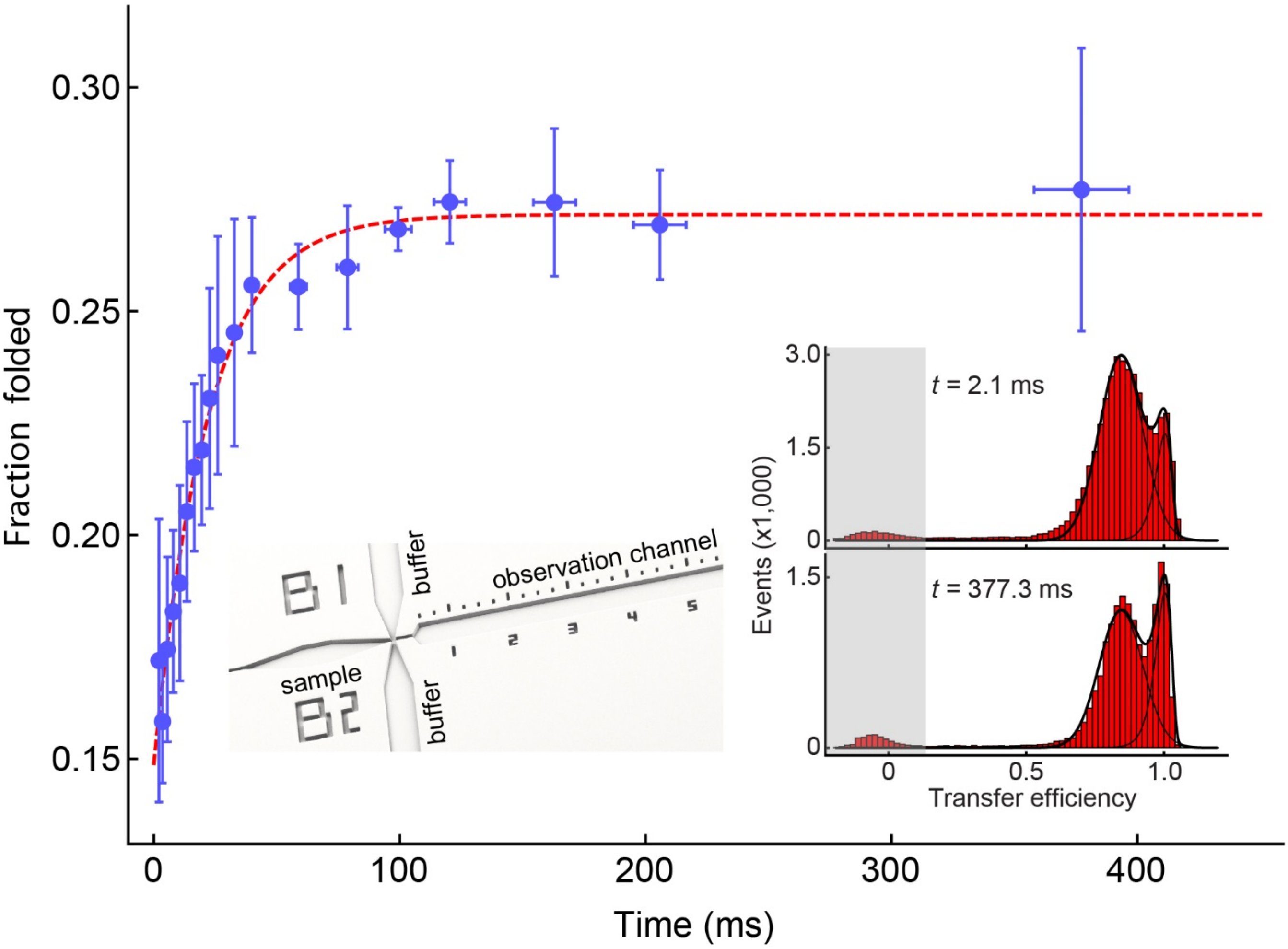
(**a**) Photograph of the temperature-controlled cell culture dish holder for the intracellular single-molecule FRET experiments. (**b**) Schematic drawing of the holder. (**c**) Temperature calibration based on the temperature-dependent fluorescence lifetime measurements of rhodamine B shows the conversion of the value measured at the sensor in the aluminum block to the actual temperature in the confocal volume. Error bars reflect the uncertainty in the fluorescence lifetime measurements used for calibration.



Supplementary Figure 8

FRET histograms of Yff1 in HBSS as a function of temperature.

FRET histograms of Yfh1 in HBSS as a function of temperature showing the heat and cold denaturation and the temperature-induced collapse of unfolded Yfh1 *in vitro*. The dashed red line indicates the position of unfolded Yfh1 at 279 K. The peak at $E \approx 0.85$ corresponds to the folded population, the peak at lower transfer efficiencies to the unfolded population of Yfh1. The peak close to a transfer efficiency of zero (shaded) corresponds to a population of molecules lacking an active acceptor chromophore. Solid lines are fits with two Gaussian peak functions (for details of the fitting procedure, see Online Methods) corresponding to unfolded and folded subpopulations (sum shown as thick solid line).



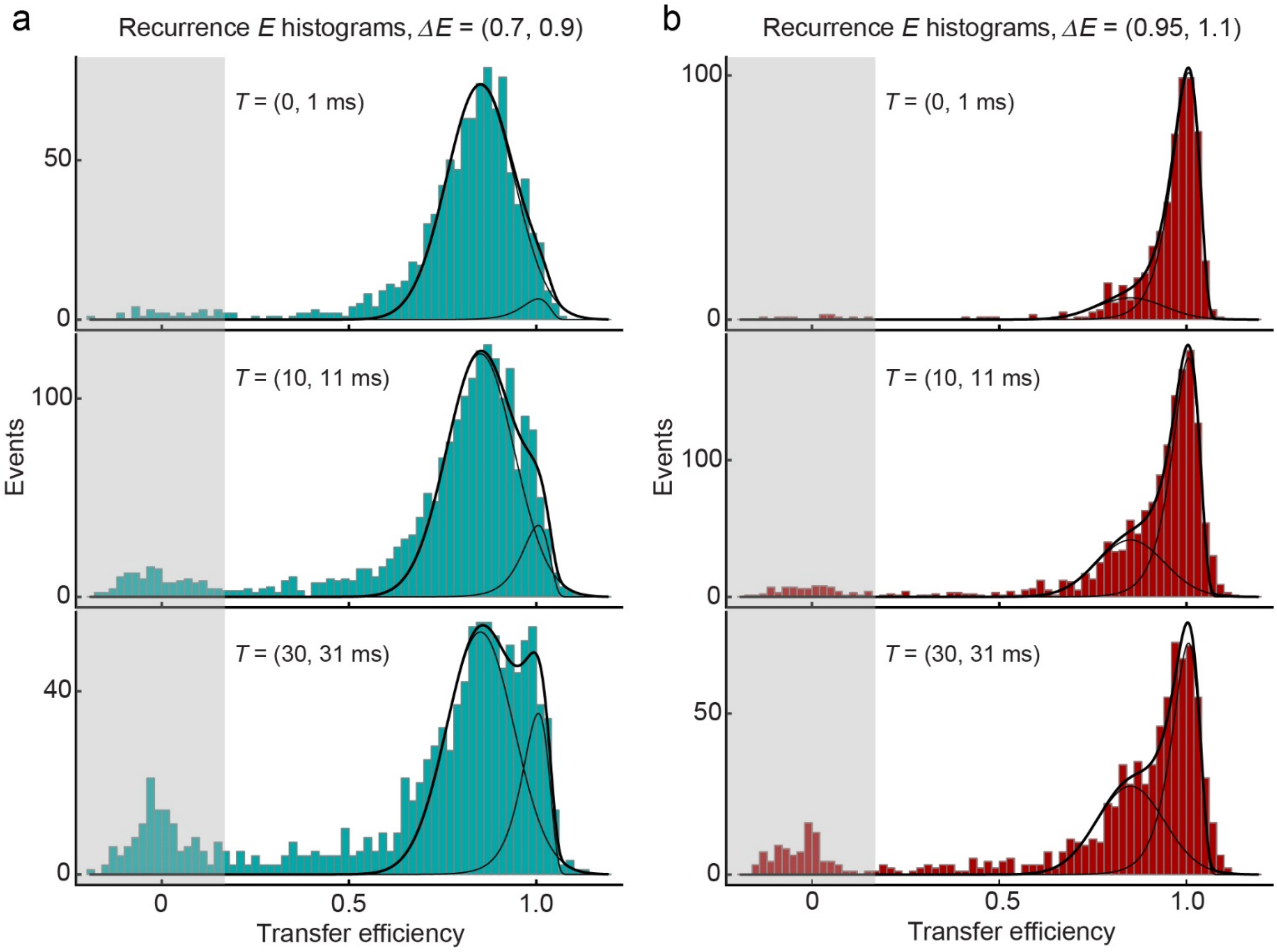
Supplementary Figure 9

Microfluidic mixing experiment of the folding of GB1 in buffer.

Microfluidic mixing experiment showing the folding of GB1 after rapid dilution of GB1 in 0.8 M GdmCl to 80 mM GdmCl at 4 °C. The error bars along the time axis describe the uncertainty originating from the positioning of the confocal volume along the observation channel and the uncertainty in flow velocity. For the uncertainty in the arrival time, we assumed a variation in the flow velocity of 5% and an uncertainty of 0.5 μm in positioning the laser focus[7]. The error bars in ordinate direction indicate the standard deviation of the fraction folded from the FRET efficiency histograms of three individual measurements. The insets show an electron microscopy image of the microfluidic mixing chamber with the observation channel, the two buffer inlets and the sample inlet channel (left), and representative FRET efficiency histograms measured after 2.1 ms and 377 ms, respectively (right). The peak close to a transfer efficiency of zero (shaded) corresponds to a population of molecules lacking an active acceptor chromophore. The fraction of unfolded GB1 ($\langle E \rangle = 0.85$) decreases over time and the fraction of folded GB1 ($\langle E \rangle = 1.00$) increases.

References:

7. Wunderlich, B. et al. Microfluidic mixer designed for performing single-molecule kinetics with confocal detection on timescales from milliseconds to minutes. *Nat. Protoc.* **8**, 1459-1474 (2013).

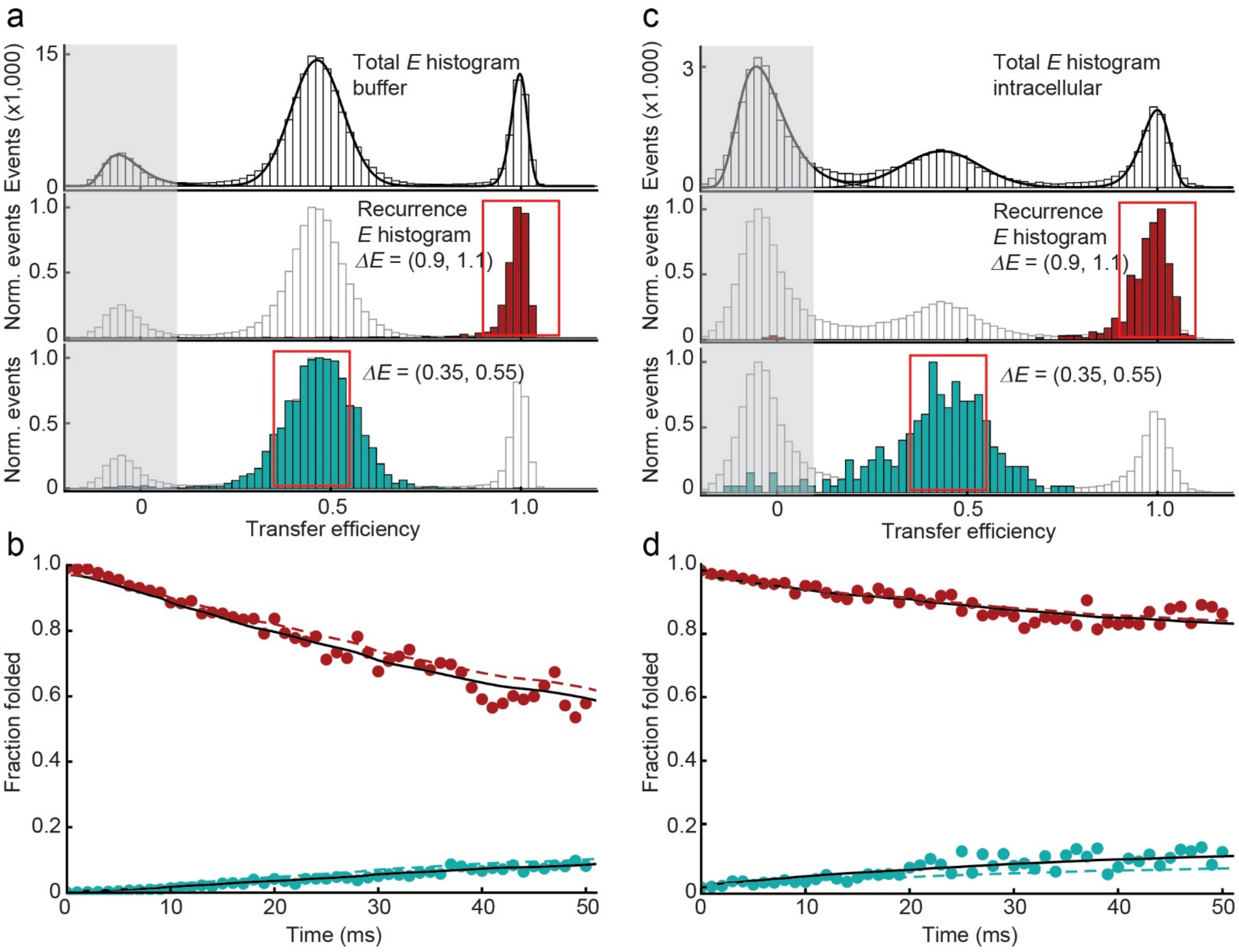


Supplementary Figure 10

Recurrence transfer efficiency histograms of GB1 in HeLa cells.

(**a**) Recurrence transfer efficiency histograms[8] of GB1 measured intracellularly with an initial transfer efficiency range *ΔE* of 0.7 – 0.9 (corresponding to the unfolded state), showing the depopulation of unfolded GB1 with increasing recurrence intervals, *T*. Starting from the unfolded state, the increase in the folded-state population at $\langle E \rangle$ = 1.00 originates from the increase in the probability of observing a folded molecule with increasing time between the first and second observed fluorescence burst due to the folding/unfolding dynamics. The peak close to a transfer efficiency of zero (shaded) corresponds to a population of molecules lacking an active acceptor chromophore. (**b**) The conversion of folded to unfolded GB1 is shown in the recurrence histograms with an initial transfer efficiency range *ΔE* of 0.95 – 1.1 (corresponding to the folded state). Starting from the folded state, the increase in the unfolded-state population at lower transfer efficiency originates from the increase in the probability of observing an unfolded molecule with increasing time between the first and second observed fluorescence burst due to the folding/unfolding dynamics. A global analysis of 61 histograms of this type was used for the kinetics shown in **Figure 4**.

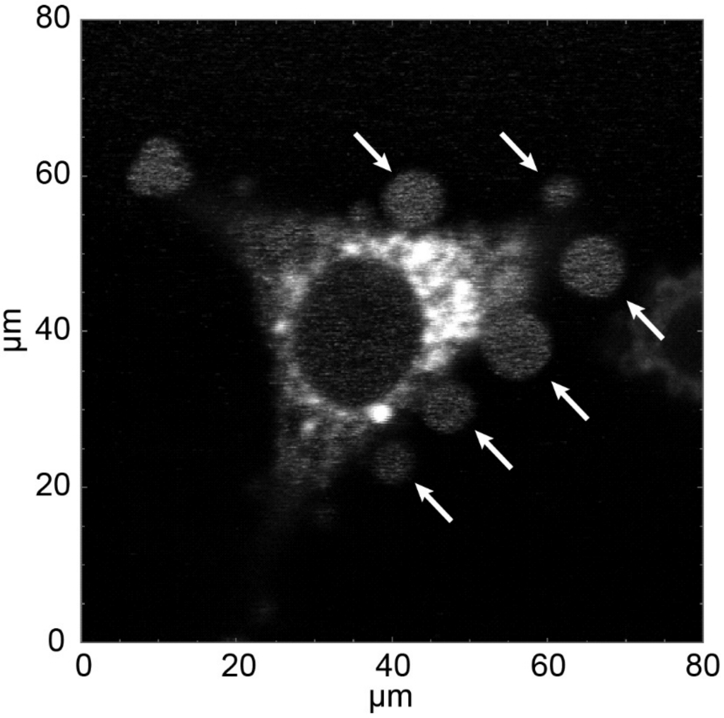
References:
8. Hoffmann, A. et al. Quantifying heterogeneity and conformational dynamics from single molecule FRET of diffusing molecules: recurrence analysis of single particles (RASP). *Phys. Chem. Chem. Phys.* **13**, 1857-1871 (2011).



Supplementary Figure 11

RASP control measurement with a noninterconverting mixture of ProTα and GB1 wild type (wt).

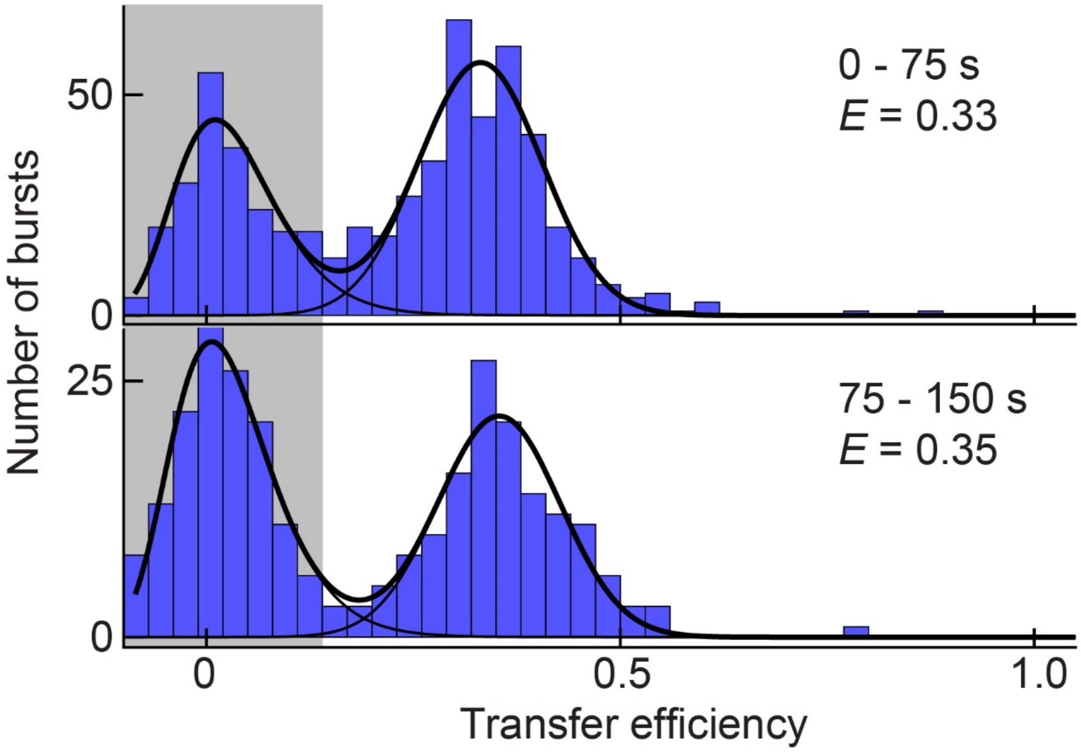
To illustrate the results from RASP[8] for two species that do not interconvert, a mixture of FRET-labeled ProTα and GB1 wt were co-injected into HeLa cells. (**a**) FRET efficiency histogram of a mixture of 20 pM ProTα and 20 pM GB1 wt in HBSS, pH 7.2, measured for 15 h at 4 °C, with recurrence histograms of GB1 wt (red, initial *E* range *ΔE* = (0.9 – 1.1)) and ProTα (cyan, initial *E* range *ΔE* = (0.35 – 0.55)), both within a recurrence interval of *T* = (0, 1 ms). The peak close to a transfer efficiency of zero (shaded) corresponds to a population of molecules lacking an active acceptor chromophore. (**b**) Recurrence analysis indicates that changes in the fraction of folded molecules with increasing delay times are only due to the occurrence of "new" (i.e. non-recurring) molecules, as shown by the coincidence between a global fit of the data with a model[8] including both folding and unfolding of recurring molecules and the appearance of new molecules (solid line) and the changes expected from the arrival of new molecules alone (dashed lines). (**c**) The same behavior was found with GB1 wt and ProTα injected into HeLa cells (red and cyan, respectively, with the same initial *E* ranges as in (**a**)). The total *E* histogram was constructed from measurements in 51 individual cells with a total measurement time of 2.8 h. (**d**) Similar to the finding in buffer, no interconversion between the two populations was observed but only changes due to the arrival of new molecules.



Supplementary Figure 12

Example of a HeLa cell showing membrane protuberances after microinjection.

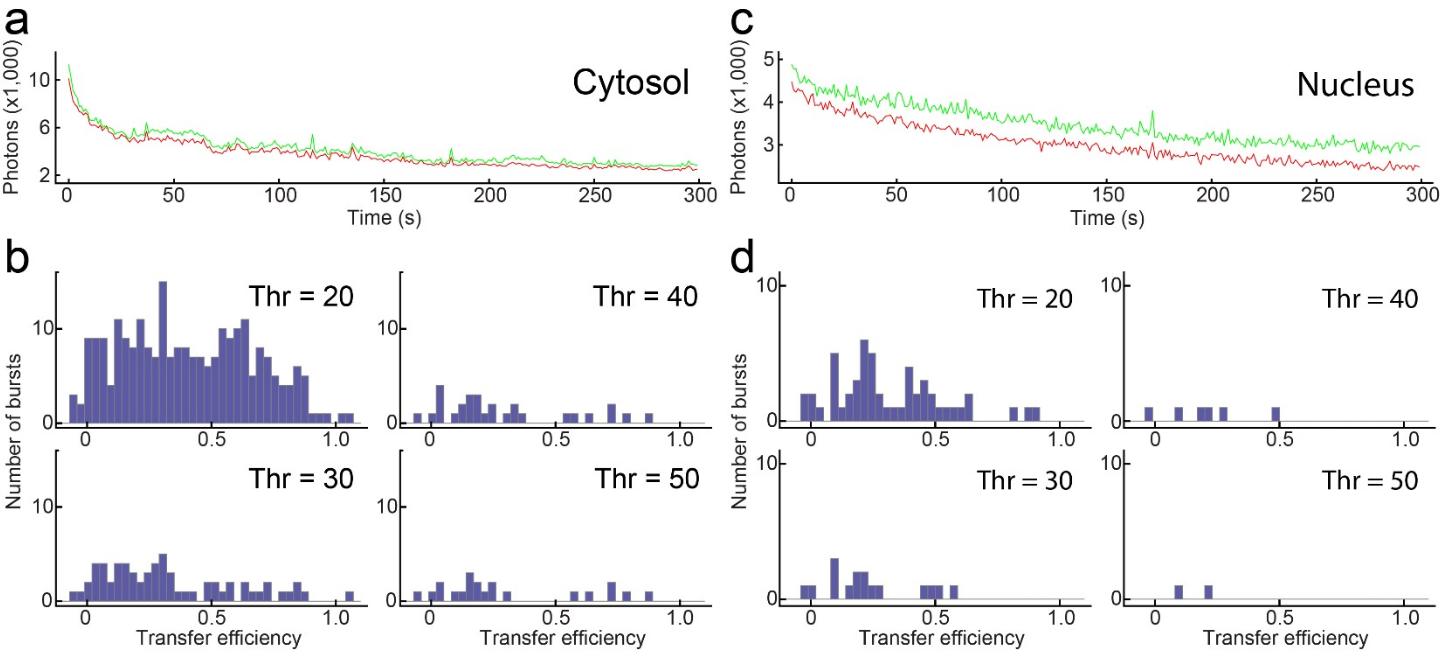
Fluorescence intensity scan of a HeLa cell microinjected using too high a pressure. One injection with a pressure of 1 bar and a pulse duration of 50 ms was applied to inject a solution of fluorescently labeled ProTα. The formation of membrane protuberances after injection (indicated with white arrows) clearly indicates that the applied pressure was too high or the pulse duration too long. Injection settings were chosen in order that the formation of such protuberances did not occur but sufficient fluorescently labeled sample was injected.



Supplementary Figure 13

Change of the FRET efficiency distribution during a measurement in a single cell.

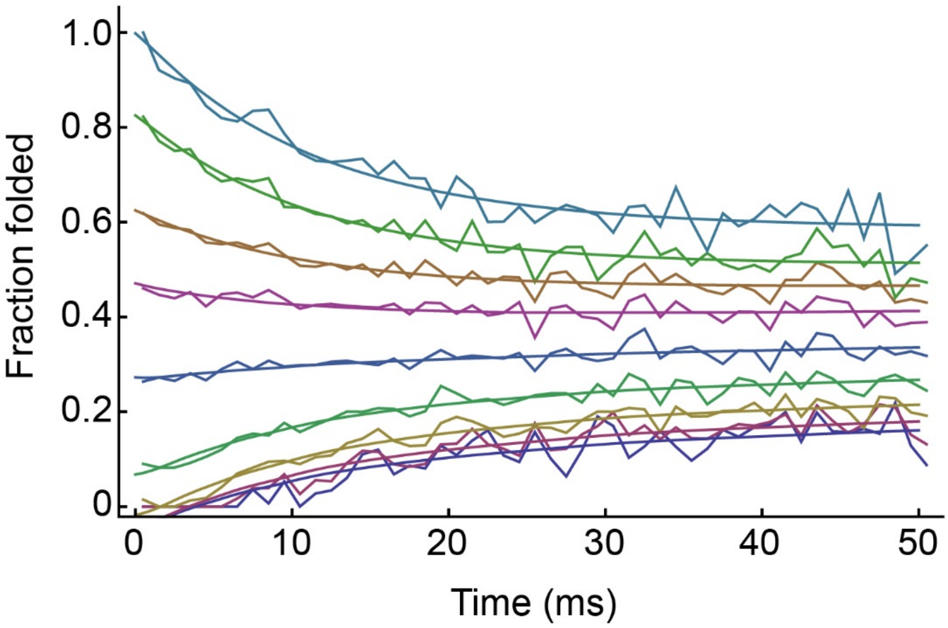
A single-molecule FRET measurement in a HeLa cell after injection of fluorescently labeled ProTα into the cytosol was split into two halves. The FRET efficiencies of doubly labeled ProTα (transfer efficiency of 0.33 to 0.35) are constant within uncertainty during the measurement in the cytosol of a single cell.



Supplementary Figure 14

Cellular autofluorescence makes a negligible contribution to FRET efficiency histograms in microinjected cells.

The contribution of cellular autofluorescence was investigated by performing single-molecule FRET measurements in non-injected cells. (**a**,**c**) Fluorescence recordings of donor (green) and acceptor emission (red) with 1-s time binning, acquired in the cytosol (**a**) and nucleus (**b**) of HeLa cells that had not been injected with fluorescently labeled protein. (**b**) FRET efficiency histograms from the measurement in the cytosol of a non-injected cell (**a**), with different thresholds (20 – 50 photons per burst) for burst detection applied. (**d**) FRET efficiency histograms from the measurement in the nucleus of a non-injected cell (**c**). In both cases, above a threshold of 30 (the minimum threshold used for injected cells), the number of bursts due to cellular autofluorescence is small and their contribution to the FRET efficiency histograms measured in injected cells is therefore negligible.



Supplementary Figure 15

Choosing different initial transfer efficiency intervals for recurrence analysis (RASP) does not affect the observed relaxation dynamics.

To illustrate that the kinetics obtained from recurrence analysis is independent of the choice of initial transfer efficiency intervals, as expected from theoretical considerations[8], we calculated the fractions of folded intracellular GB1 from recurrence FRET efficiency histograms analogously to the curves shown in **Figure 4f** but with a broad range of different initial transfer efficiency intervals, from bottom to top: *ΔE* = (0.6, 0.8), (0.65, 0.85), (0.7, 0.9), (0.75, 0.95), (0.8, 1.0), (0.85, 1.05), (0.9, 1.1), (0.95, 1.15), and (1.0, 1.2). All data (noisy curves) can be described with one relaxation time of 12 ms (smooth lines), the value we obtained from the analysis of the data shown in **Figure 4f**, where the intervals were *ΔE* = (0.7, 0.9) and *ΔE* = (0.95,1.1).

Supplementary Table 1

Amino acid sequences of prothymosin alpha (ProTα), the IGG-binding domain of protein G (GB1), GB1 wildtype (GB1 wt), and the yeast frataxin homolog Yhf1.

The labeling positions are indicated in underlined boldface.

		1	11	21	31	41
ProTα	GP	**C**DAAVDTSSE	ITTKDLKEKK	EVVEEAENGR	DAPANGNAEN	EENGEQEADN
(S1C/E56C)		51	61	71	81	91
		EVDEE**C**EEGG	EEEEEEEEGD	GEEEDGDEDE	EAESATGKRA	AEDDEDDDVD
		101				
		TKKQKTDEDD				
		1	11	21	31	41
Yhf1		MESSTDGQVV	PQEVL**C**LPLE	KYHEEADDYL	DHLLDSLEEL	SEAHPDCIPD
(V1M/N16C/ S120C)		51	61	71	81	91
		VELSHGVMTL	EIPAFGTYVI	NKQPPNKQIW	LASPLSGPNR	FDLLNGEWVS
		101	111	121		
		LRNGTKLTDI	LTEEVEKAI**C**	KSQ		
		1	11	21	31	41
GB1	GAM	G**C**YKLILNGK	TLKGETTTEA	VDAATAEKVL	KQYANDNGVD	AEWTYDDATK
(Q2C/F30L/ G41A/T55C)		51				
		TFTV**C**E				
		1	11	21	31	41
GB1 wt	GAM	G**C**YKLILNGK	TLKGETTTEA	VDAATAEKVF	KQYANDNGVD	GEWTYDDATK
(Q2C/T55C)		51				
		TFTV**C**E				

4 Protocol for microinjection of fluorescently labeled proteins into adherent mammalian cells for single-molecule fluorescence spectroscopy

Iwo König & Benjamin Schuler

Protocol for microinjection of fluorescently labeled proteins into adherent mammalian cells for single-molecule FRET spectroscopy

Iwo König[1] & Ben Schuler[1]

[1]Department of Biochemistry, University of Zurich, Zurich, Switzerland

Abstract

We have recently shown that microinjection is a reliable and efficient way to deliver fluorescently labeled proteins into living mammalian cells for single-molecule FRET spectroscopy. The following protocol gives a detailed description of how fluorescently labeled proteins are microinjected into HeLa cells, how in-cell data are recorded, and how they are analyzed (Fig. 1).

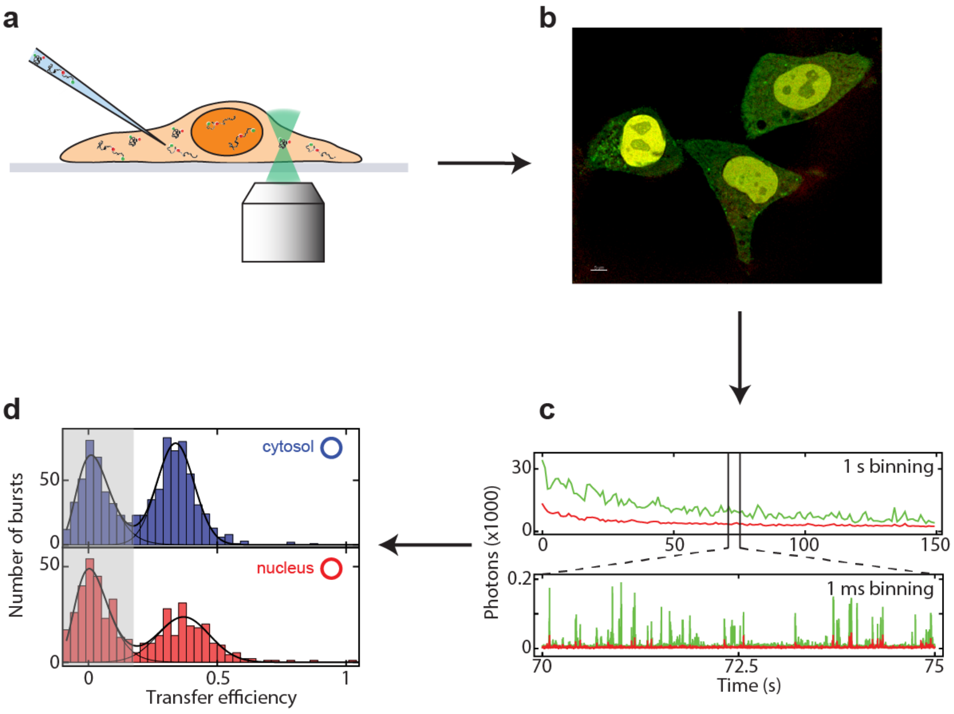


Fig. 1: Overview over the process of measuring single-molecule FRET in live cells. The process consists of microinjection (a), analyzing injected cells by fluorescence imaging (b), measuring single-molecule fluorescence data (c), and analyzing the data to obtain e.g. FRET histograms (d).

Introduction

Microinjection is a powerful tool to deliver suitable amounts of fluorescently labeled proteins into living mammalian cells to perform single-molecule fluorescence measurements. Here, we provide a step-by-step protocol for microinjecting and measuring FRET-labeled proteins in adherent mammalian cells. This protocol provides a more detailed description on the basis of the recently published methods paper (1), with the aim to perform single-molecule FRET, FCS, and nanosecond-FCS (nsFCS) measurements in living cells. The main steps consist of cultivating adherent mammalian cells, preparing the sample and the microinjection devices, performing microinjections and single-molecule FRET measurements, and the data analysis. It is assumed that the experimenter is aware of how to operate a confocal single-molecule fluorescence microscope and analyze the data, and that a suitable purified sample labeled with fluorescent dyes is available (2).

Overview of the procedure

1) Cultivation of cells and seeding into microscopy dishes (1 h, overnight incubation)
2) Preparation of the microscope and the microinjection device (Fig. 2) (0.5 h)
3) Sample and microinjection capillary preparation (1 h)
4) Sample injection and in-cell measurements (6 – 10 h)
5) Data analysis (~5 h)

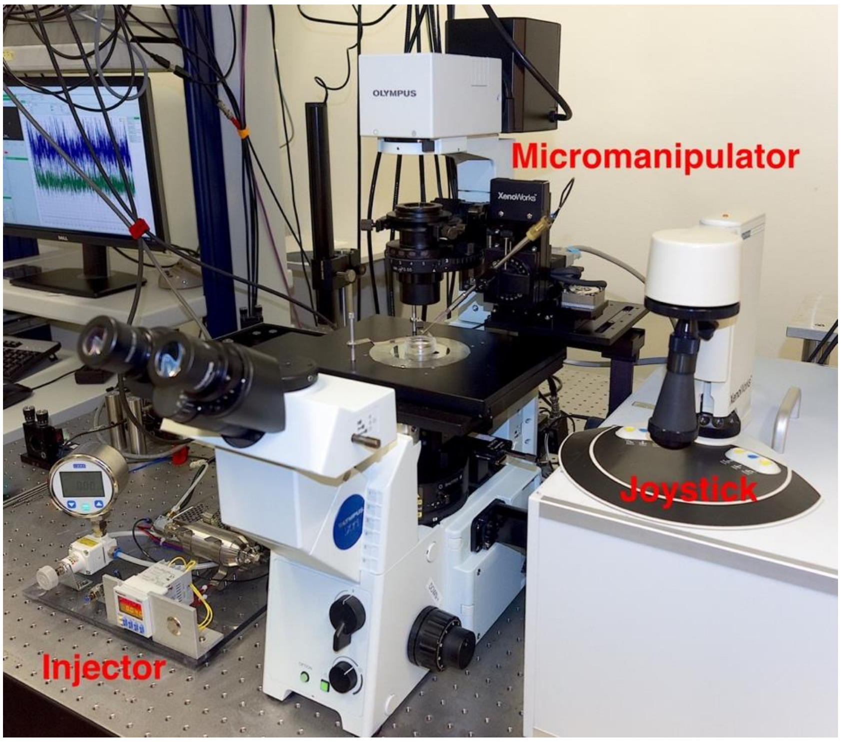


Fig. 2: The complete assembly for microinjection consisting of the inverted microscope, a pneumatic injector, and a micromanipulator with joystick for controlling the position of the capillary.

Materials

Fluorescently labeled protein

- Site-specifically labeled protein, with a dye-pair suitable for FRET and excitation wavelengths at or above ~520 nm, e.g. Atto 532 or Cy3B as a donor in combination with Biotium CF680R, Abberior Star 635, or Atto 647N as an acceptor. **Comment 1**

Cell culture

- Adherent mammalian cells such as HeLa cells (American Type Culture Collection). **Comment 2**
- 75 cm^2 tissue culture flasks (TPP)
- A suitable cell culture medium without phenol red, e.g. Dulbecco's Modified Eagle Medium (DMEM, Invitrogen), supplemented with 10% fetal calf serum (FCS, Bioconcept Amimed), 100 units/ml penicillin, and 0.1 mg/ml streptomycin (Sigma)
- Phosphate buffered saline (PBS, Sigma)
- Hanks' balanced salt solution (HBSS, Sigma)
- 0.5 g/l trypsin and 0.2 g/l EDTA solution (Sigma) for cell detachment
- 35-mm glass-bottom microscopy dishes (Ibidi)
- 0.01% poly(l-lysine) solution (PLL, Sigma)

Instrumentation for microinjections

- Xenoworks (Sutter Instrument) or similar microinjection system
- Custom-built microinjector or suitable commercial system, e.g. Picospritzer III (Parker) or Femtojet (Eppendorf)
- Femtotips II (Eppendorf) or equivalent microinjection capillaries
- Microloader tips (Eppendorf)
- 1 mg/ml poly(l-lysine)-graft-poly(ethylene glycol)57 (PLL(20)-g[3.5]-PEG(2), SuSoS)
- 0.1% Tween 20 (Pierce)

Instrumentation for in-cell FRET/FCS measurements

- Confocal single-molecule instrumentation, e.g. MicroTime 200 (PicoQuant), using an inverted microscope body (e.g. Olympus IX71 with an Olympus UplanApo 60×/1.20W objective)
- Suitable excitation source, e.g. 20-MHz supercontinuum laser (SC-450-4, Fianium) in combination with a 520/15 band-pass filter (Chroma Technology) or 532-nm continuous-wave laser (LaserBoxx LMX-532S, Oxxius)
- LP532 long-pass filter (Chroma Technology)
- 100-µm confocal pinhole
- Polarizing beam splitter
- 635DCXR dichroic mirrors (Chroma Technology)
- ET585/65m donor filters (Chroma Technology)
- LP647RU and HC750/SP acceptor filters (Chroma Technology)

- τ-SPAD avalanche photodiodes (PicoQuant) or SPCM-AQR-15 single-photon avalanche diodes (PerkinElmer).
- HydraHarp 400 counting module (PicoQuant)
- Piezo stage combination (P-733.2 and PIFOC, Physik Instrumente GmbH) or equivalent type of 3D scanning

Procedure

Cultivating cells

1) Keep cells in culture in 75 cm^2 tissue culture flasks in humid atmosphere at 37 °C and 5% CO_2, using phenol-red free culture medium supplemented with FCS and penicillin/streptomycin.
   Caution: Test cell cultures for the presence of mycoplasma, e.g. using the MycoAlert mycoplasma detection kit (Lonza)!
2) Subculture cells when approximately 80% confluency is reached using trypsin-EDTA for cell detachment.

Preparation of cells for microinjection

3) Clean as many 35-mm glass-bottom microscopy dishes as needed by rinsing them twice with sterile PBS to remove impurities on the glass surface.
4) Seed ~2×10^5 cells in 4 ml phenol red–free culture medium into a 35-mm glass-bottom microscopy dish and incubate overnight (37 °C, 5% CO_2).
   Caution: 100 % confluency results in fast cell detachment during injections. The best cell adherence was achieved at 50 – 70 % confluency.
   Caution: To improve cell adherence, the glass surface of the microscopy dishes can be pre-coated with 0.01% PLL solution for 10 min, followed by two washing steps with sterile H_2O.
5) Prior to single-molecule measurements, aspirate the medium and wash the cells twice with the buffer of choice (measurement buffer).
6) Cover the cells with 3 ml measurement buffer.
   Critical step: Choose a buffer that supports the cells with nutrients for the duration of the experiment. However, the buffer should not contribute to the fluorescence background (e.g. originating from phenol red or fetal calf serum), and the buffer components should not interact with the sample molecules (e.g. conformational changes of the protein upon binding of Ca^{2+} or Mg^{2+} present in the buffer). For longer-term measurements, heat the sample environment to 37 °C and, if necessary (especially for carbonate-based buffer), provide 5% CO_2 atmosphere to avoid changes in pH.

Preparation of sample and capillaries for microinjection

7) Dilute labeled protein into measurement buffer (+ 0.001 % Tween 20 to minimize surface adhesion in capillary) to a final concentration of ~20 – 100 nM.
   Critical step: Choosing a suitable concentration depends on the properties of the protein and needs to be optimized for every protein individually. Single-molecule FRET measurements require ~20 – 50 nM initial sample concentrations, nsFCS measurements ~50 – 100 nM.
8) Dispense 5 µl of sample into a FemtoTip II capillary using Microloader tips.
   Caution: If you work with proteins that stick to glass surfaces, prior coating of the FemtoTip capillary might be required. Incubate the capillary with 10 µl of a 1 mg/ml PLL-PEG solution for 10 min, followed by rinsing with H_2O.
9) Mount the loaded capillary on the micromanipulator (Fig. 3).

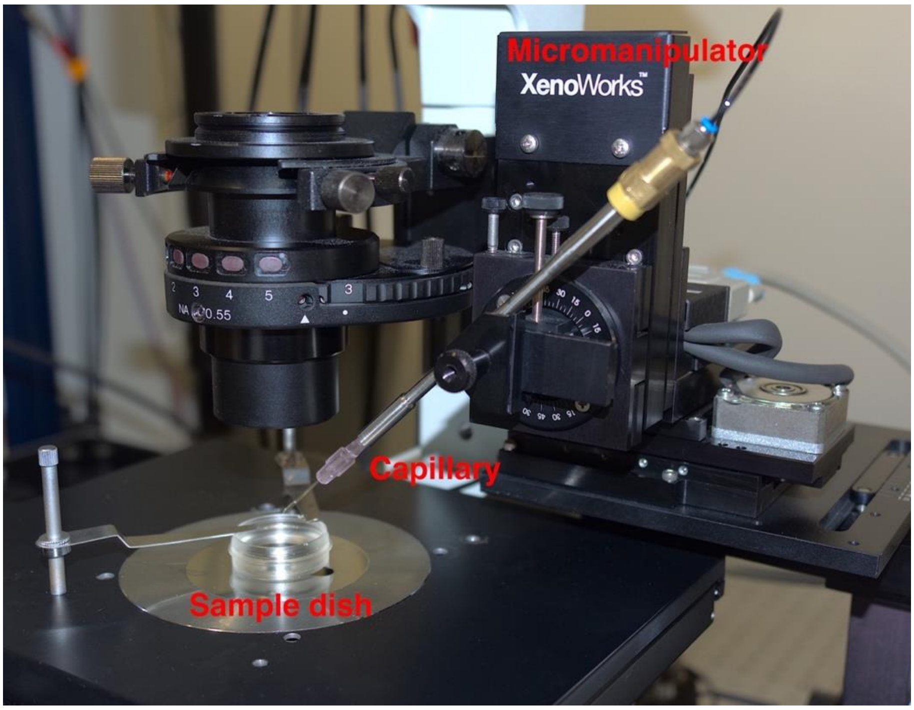


Fig. 3: Micromanipulator with mounted capillary and glass-bottom microscopy dish. A narrow light beam from the condenser can be used to aid positioning the tip of the capillary and localizing the needle with the binocular eyepiece.

Testing the integrity of the injection capillary and the sample ('stream test')

10) Place a microscopy dish, filled only with measurement buffer, on the microscope and focus with the objective to approximately 5 – 10 µm above the glass surface.
11) Position the tip of the capillary above the objective lens. A narrow light beam from the condenser, with the aperture maximally closed, can be used as a guide.
12) Immerse the tip of the capillary in the measurement buffer with the micromanipulator and use the binocular to find the shadow of the capillary. When the capillary is in the field of view, move it downwards until the tip is close to the glass surface, which is when the tip gets into focus.
    Caution: Use the joystick in a coarse gear to find the capillary, to immerse it into the measurement buffer, and to move it into the field of view. Switch to a finer gear for moving the tip into focus to avoid breaking it by collision with the glass surface.
13) Set the pressure of the injector to 2-3 bar and the pulse duration to 50 ms (Fig. 4).
14) Inject several times into the measurement buffer and check whether the capillary releases sample into the measurement buffer. Indications for an intact capillary can be, e.g., the ejection of an initial air bubble from the needle or a shadow of the stream due to a refractive index mismatch with the surrounding medium.
    Critical step: The tip of the capillary is easily blocked. To unblock it, gently

move the tip of the capillary across the glass surface while constantly pushing the inject button at the injector until sample starts to flow.

15) Use the crosshair in the binocular eyepiece to position the tip of the needle to the side of the confocal volume to ensure that the sample stream will be ejected towards the confocal observation volume.
Caution: The needle should be positioned sufficiently far away from the focus (~50 µm) to ensure that the flow does not influence the diffusion time of the labeled molecules determined by FCS.

16) Reduce the injection pressure to 0.5 bar.

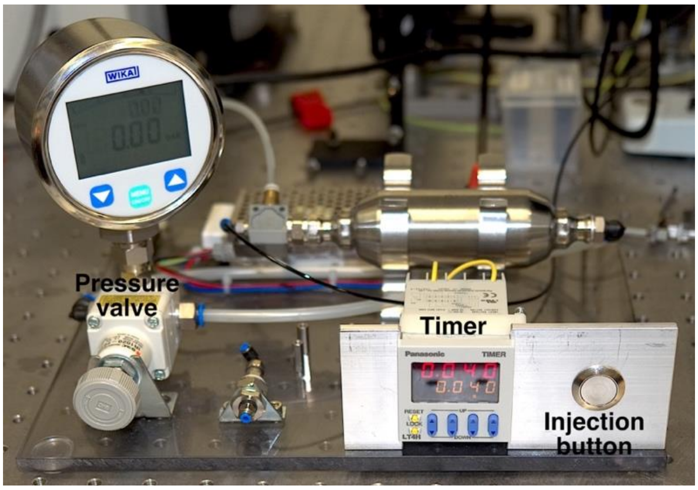


Fig. 4: The custom-built injector. The pressure of the pulse can be set with the valve on the left side. The timer at the right allows the duration of the pressure pulses for injections to be set. Analogous settings are available in commercial injection systems.

17) Set the intensity of the excitation laser to 50 µW, measured at the back-aperture of the microscope, and open the detectors. **Comment 3**

18) Start recording data and inject sample several times into the measurement buffer until fluorescence bursts are detected.

19) Continue injecting for a total of ~3 min. Inject with a low pressure (0.5 bar) to maintain a frequency of bursts in the single-molecule regime (i.e., baseline-separated bursts).

20) Analyze the resulting transfer efficiency histogram to check the integrity of the sample (Fig. 5). **Comment 4**

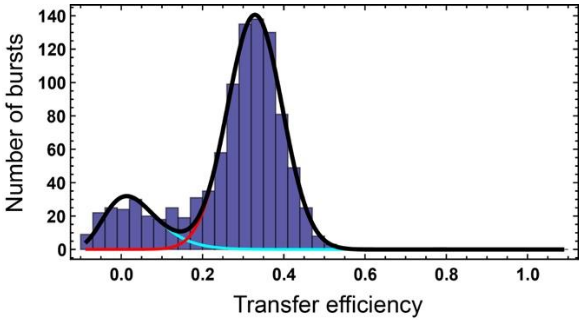


Fig. 5: Example of a FRET efficiency histogram obtained from a stream test, where fluorescence bursts were recorded for 3 min. The acceptor-inactive population ($\langle E\rangle \cong 0$) and the doubly labeled population ($\langle E\rangle \cong 0.35$) were fitted with a log-normal and a Gaussian peak function, respectively.

Microinjection into cells

21) Mount a microscopy dish with adherent cells and move the injection capillary into focus (steps 10 – 12).
22) Set the injection pressure to 0.5 bar and the pulse timer to 50 ms.
23) Inject several times into the buffer to rinse the capillary.
Caution: Rinsing is needed because of a steady flux of measurement buffer into the needle due to capillary forces or hydrostatic pressure.
24) Move the tip of the capillary to a position above the cytosol or nucleus of a cell.
25) Move the capillary downwards until the cell membrane starts to deform visibly.
26) Inject once by triggering the injector and observe the cell. If the cell inflates during the pressure pulse and deflates rapidly afterwards, injection was successful. Inject ~3 – 5 cells at once to optimize the experimental throughput.
Caution: If spherical extrusions of the cell membrane (or even cell rupture) are observed after injection, the pressure and/or the duration of the injection pulse must be decreased until full recovery of the cell shape after injection is attained. Once the injection settings are optimized for a given sample and cell type, they can usually be used for the entire series of injections.

Single-molecule FRET/FCS measurements

27) For choosing the optimal position of the confocal detection volume in the cell, first scan the injected cell in the XZ or YZ plane using the Piezo stage (fluorescence lifetime (FLIM) or fluorescence intensity image). Then set the axial (Z) position of the focus ~1 – 3 μm above the glass surface, but still within the cell.
Caution: Switch on the laser only during scans or measurements to reduce bleaching of the injected protein. It is usually sufficient to scan the cells with reduced laser power, e.g. using an OD 1 filter.
28) Subsequently perform an XY scan of the injected cells to identify and set the lateral (XY) position of the focus within the cell. **Comment 5**
Caution: Do not choose very bright regions in the cell for a measurement, since cellular background is likely to dominate the overall signal in these areas (fluorescence of injected sample and cellular components can often be distinguished by FLIM (1)); additionally, there is the risk of saturating the detectors.
29) Acquire data for ~3 min at the position chosen.
Caution: If there are still fluorescent bursts detected after 3 min, the measurement can be continued until most of the injected protein molecules are photobleached.

Data analysis – FRET efficiency histograms

In principle, in-cell data are analyzed in the same way as regular measurements on freely diffusing molecules in buffer solution, but one should consider the following points:

30) Split the time trace into short segments where the background can be assumed to be approximately constant (Fig. 6).

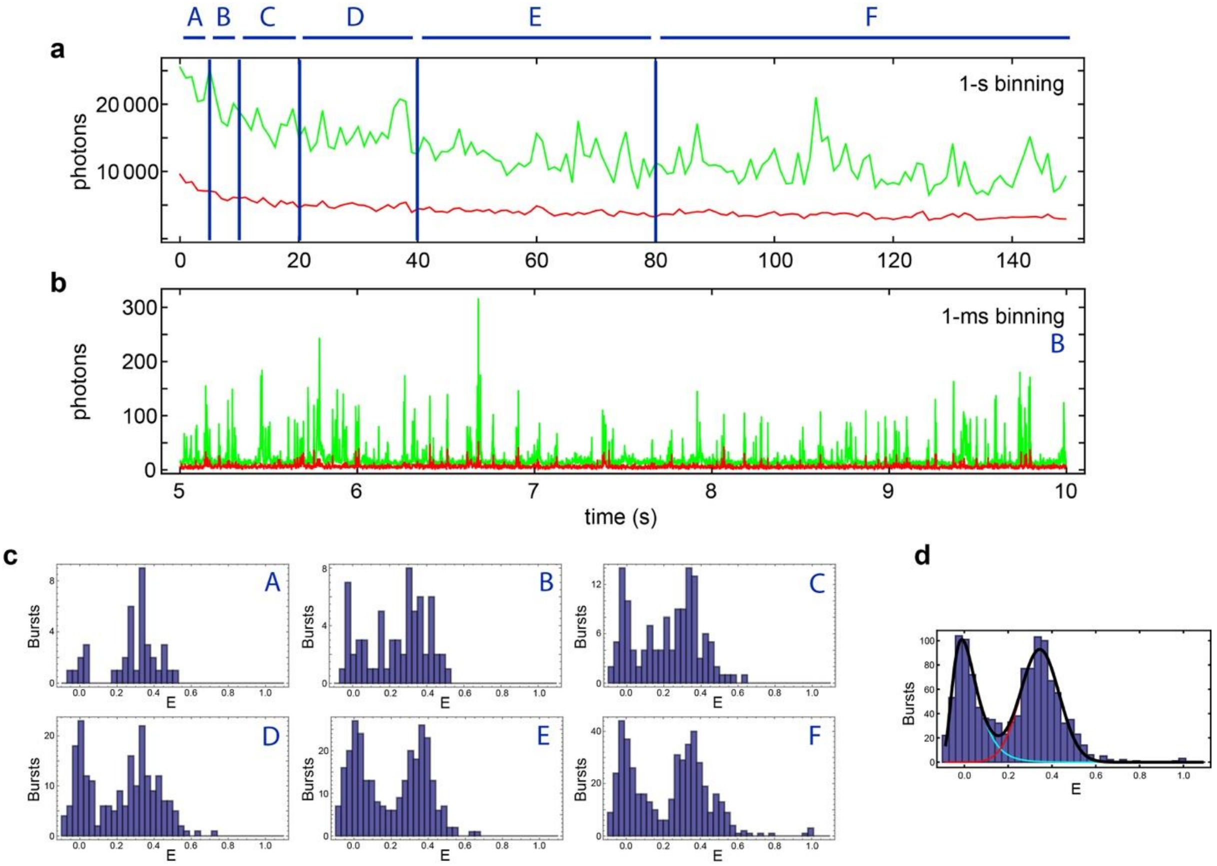


Fig. 6: Example of an intracellular single-molecule FRET measurement after injection of labeled proteins. (a) Fluorescence recording of a cytosolic single-molecule FRET measurement with 1-s time binning. The trace is split into segments A – F at 5 s, 10 s, 20 s, 40 s, and 80 s to create sections with almost constant background. (b) Blow-up of the fluorescence recording of segment B (5 – 10 s) with 1-ms binning showing background-separated fluorescence bursts. (c) FRET efficiency histograms of segments A – F. (d) Histograms A – F were summed up to obtain the final FRET efficiency histogram from one single cell. The populations were fitted according to **Comment 5**.

31) Identify fluorescence bursts in every segment by binning the recording with a bin with of 1 ms and using a threshold of ~40 to ~100 photons/bin after background (average signal without detected bursts) subtraction, depending on the intensity of the background signal in the respective section, and combine contiguous identified bins.
32) Build a transfer efficiency histogram from the combined bursts of the individual segments.
33) Analyze the histogram as required, e.g. by fitting with suitable peak functions for individual subpopulations. **Comment 6**

Data analysis – FCS

34) Correlate the data from the stream test of the capillary (1) to a minimum time of ~10^{-5} s and fit the donor-acceptor cross-correlation (which is specific for FRET-labeled molecules) to obtain the diffusion time in buffer.
35) Correlate the measured intracellular data and analyze the donor-acceptor cross-correlations accordingly to obtain the intracellular diffusion times. The comparison of the correlation curves and the translational diffusion times enables an assessment of diffusion in the cell versus the medium and

provides an additional cross-check that the confocal volume was correctly positioned inside the cell.

In-cell nsFCS measurements

36) Load the capillary with a sample containing ~50 to ~100 nM labeled protein and inject cells as described above (steps 7 – 9, 21 – 26).
37) Scan the injected cells with a low laser power to locate the injected cells and position the laser focus (steps 28 - 29).
38) Change to T2 mode and unplug the sync cable.
39) Record data for 2-3 min and repeat in other injected cells until sufficient statistics are available (at least a total measurement time of ~2 h, e.g. 40 – 50 individual cells).

Data analysis – nsFCS

40) Correlate the measured data as described earlier (1, 3, 4).
41) Average only data from cells where the correlation curve shows a good signal-to-noise ratio (indicated e.g. by the visibility of the antibunching component at $\tau = 0$ ns).

Comments

1) Fluorescent dyes for in-cell single-molecule experiments should be chosen according to their availability for site-specific labeling (e.g. maleimide or succinimidyl chemistry) and their photophysical (photostability and brightness) and chemical (hydrophobicity) properties. Ideally, the dyes should not interact attractively with the protein or with the intracellular environment, which would reduce the rotational mobility of the dyes on the protein and affect protein stability and/or the translational mobility of the sample in the cell. Especially very hydrophobic dyes such as Atto 647N, which binds to mitochondria (5), should be avoided or need to be tested carefully for each individual case. So far, dye pairs such as Atto 523/Biotium CF680R or Cy3B/Biotium CF660R showed good performance in terms of photophysical properties and the absence of unwanted interactions with the sample or the intracellular environment.
2) Any type of adherent cells suitable for microinjection can be used, ideally growing in a culture medium without phenol red. If long-term growth in media without phenol red is not feasible, the cells should be washed 3 times with a medium without phenol red before the experiment to reduce background fluorescence.
3) Either pulsed or continuous wave (cw) excitation can be used for in-cell experiments. Pulsed excitation has the advantage that fluorescence lifetimes can be obtained from every measurement. Also, the combination with a red laser for acceptor excitation enables alternating excitation of the donor and the acceptor dye (e.g. PIE (6) or ALEX (7)). However, with pulsed excitation, especially the acceptor fluorophores tend to photobleach faster than with cw excitation, resulting in an increased donor-only population and less FRET-

active molecules. Cw excitation can thus be preferable to minimize photobleaching and obtain brighter fluorescence bursts.

4) Interactions of FRET-labeled proteins with glass surfaces can lead to an inactivation of the acceptor fluorophore (for currently unknown reasons). It can thus happen that initially no acceptor photons are detected during the stream test. This complication can be overcome by either injecting more sample into the measurement buffer until acceptor photons appear, or by coating the capillary surface, e.g. with PLL-PEG, to further reduce surface interactions of the sample.
5) Immobile FRET-labeled molecules can be detected and localized by scanning injected cells (1). However, the subsequent analysis of the fluorescence intensity or fluorescence lifetime images (FLIM), to obtain, e.g., FRET efficiencies, requires that the spatially heterogeneous contribution of fluorescence background can be corrected accurately.
6) If necessary, fit parameters, e.g. the peak width, can be fixed or constrained to fit the histogram appropriately. The transfer efficiency histograms of ProTα in cells shown here were fitted using a log-normal peak function for the donor-only population with a fixed width (0.15) and asymmetry (1.32) and a Gaussian peak function for the doubly labeled population, without any constraints (1). These values were chosen to best describe the observed shape of the populations for the entire set of single-cell histograms.

References

1. König I, et al. (2015) Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells. *Nat Methods* 12(8):773–779.

2. Schuler B, Müller-Späth S, Soranno A, Nettels D (2012) Application of confocal single-molecule FRET to intrinsically disordered proteins. *Method Mol Biol* 896(Chapter 2):21–45.

3. Hillger F, et al. (2008) Probing protein-chaperone interactions with single-molecule fluorescence spectroscopy. *Angew Chem Int Ed* 47(33):6184–6188.

4. Nettels D, Gopich IV, Hoffmann A, Schuler B (2007) Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc Natl Acad Sci USA* 104(8):2655–2660.

5. Pellett PA, et al. (2011) Two-color STED microscopy in living cells. *Biomed Opt Express* 2(8):2364–2371.

6. Müller BK, Zaychikov E, Bräuchle C, Lamb DC (2005) Pulsed interleaved excitation. *Biophys J* 89(5):3508–3522.

7. Kapanidis AN, et al. (2005) Alternating-laser excitation of single molecules. *Acc Chem Res* 38(7):523–533.

5 Impact of in-cell crowding on the conformations and dynamics of an intrinsically disordered protein

Iwo König, Andrea Soranno, Daniel Nettels & Benjamin Schuler

Impact of in-cell crowding on the conformations and dynamics of an intrinsically disordered protein

Iwo König[1], Andrea Soranno[1,2], Daniel Nettels[1] & Benjamin Schuler[1]

[1]Department of Biochemistry, University of Zurich, Zurich, Switzerland
[2]Department of Biochemistry and Molecular Biophysics, Washington University, St. Louis, USA
Correspondence should be addressed to B.S. (schuler@bioc.uzh.ch).

Abstract

The interior of a cell is crowded with macromolecules such as proteins and nucleic acids. It is expected that the presence of high concentrations of crowders influence protein stabilities and conformations. Intrinsically disordered proteins (IDPs) are particularly susceptible to changes in the environment because of their lack of a folded structure and their correspondingly flat free energy surfaces. We have recently shown that IDPs adopt more compact dimensions in highly crowded solutions *in vitro*. However, no such effects were observed in live mammalian cells under usual cell culture conditions. To vary the degree of crowding, we thus reduced the cell volume by osmotic shock and used the IDP prothymosin α (ProTα) as a crowding sensor. In the cytosol of crowded cells, ProTα indeed exhibited a slight compaction, and its translational diffusion was much slower than in cells under physiological conditions. Its chain reconfiguration dynamics, however, were similar in both cases. These results, combined with *in vitro* experiments using artificial crowders and cellular protein concentration measurements, allowed us to conclude that the cytosol of commonly used eukaryotic cell lines is less crowded than commonly assumed and enabled us to quantify the length-scale dependent influence of the intracellular environment on the dynamics and dimensions of ProTα.

Introduction

Most of what we know about protein conformations and dynamics was discovered by studies carried out under chemically defined conditions in the test tube. However, there is a great interest in the effect of the intracellular environment on proteins and other biomolecules. The inside of a cell is commonly assumed to be highly crowded with metabolites and macromolecules such as proteins, nucleic acids or polysaccharides (1, 2). Previous *in vitro* experiments and simulations have shown that especially crowding due to high concentrations of macromolecules can affect protein conformations, dynamics or folding kinetics (3-6). Studying these effects directly in living cells is difficult because of the complexity of the cellular environment. Several methods have been established to observe proteins in living cells, such as fluorescence recovery after photobleaching (FRAP) (7), ensemble Förster resonance energy transfer (FRET) measurements (8) or nuclear magnetic resonance (NMR) spectroscopy (9). These methods, however, average over an ensemble of many proteins and cells. Recent developments in single-molecule fluorescence spectroscopy have enabled the investigation of single proteins in bacteria (10) and in eukaryotic cells (11-13).

Our recently developed approach to perform single-molecule fluorescence spectroscopy in living mammalian cells (13) has allowed us to determine dimensions and dynamics of intrinsically disordered proteins (IDPs) and other proteins by performing FRET measurements and fluorescence correlation spectroscopy (FCS) in the cytosol and the nuclei of adherent mammalian cells. We found that the intracellular environment had a surprisingly small effect on the dimensions and dynamics of the small acidic IDP prothymosin α (ProTα), on the folding kinetics of the IgG-binding domain of protein G (GB1), and on the conformational stability of frataxin (13). These findings were in contrast to the generally invoked picture of a heavily crowded cellular environment (14). A broad range of intracellular concentrations of macromolecules were reported in the past, from 300 - 400 mg/ml in *E. coli* (1) to 50 – 200 mg/ml in eukaryotic cells (2, 15). It is thought that macromolecular concentrations in this range can have a pronounced influence on protein stabilities and dimensions of IDPs and unfolded proteins (16). Several research groups have tried to address this question experimentally. Boersma *et al.* (17), e.g., investigated an alpha-helical hinge protein, tagged with fluorescent proteins, which showed significant compaction in *E. coli* and HEK cells. On the other hand, no significant influence of the cytosol of HeLa cells on the dimensions of fluorescently labeled and injected polyethylene glycol (PEG) was observed by Gnutt *et al.* (18). Other studies reported a dominant role of attractive interactions between cytosolic macromolecules and the sample for modulating the stabilities of different proteins, whereas the contribution of excluded volume effects was negligible (19-21).

IDPs are expected to be particularly susceptible to conformational changes induced by crowding or other intracellular interactions, owing to their flat and comparatively featureless free energy surfaces (22). It is thus surprising that no such crowding effects were observed for ProTα in live eukaryotic cells (13), even though previous observations in vitro had revealed a pronounced compaction of ProTα in crowding experiments with PEG (23). This discrepancy raises the question of whether intracellular crowding is less pronounced than expected and how cellular effects compare to in vitro experiments with artificial crowders. Here, we use the intrinsically disordered protein ProTα as a FRET-sensor for intracellular crowding by employing single-molecule fluorescence techniques in live mammalian cells. Owing to its large negative net charge of -44, ProTα is very expanded and attractive interactions with cellular components were found to be negligible (13), making it an ideal molecule for probing nonspecific excluded volume effects. By measuring dimensions and the diffusional behavior of the IDP in cells at physiological conditions and in osmotically contracted cells, we tried to shed light on the question of how crowded the inside of a eukaryotic cell actually is.

Results and Discussion

ProTα was site-specifically labeled with the fluorescent dyes Atto 532 and Biotium CF680R and microinjected into live HeLa cells (Fig. 1a). Under physiological conditions, ProTα diffused freely through the cytosol and nucleus of injected cells. ProTα has a C-terminal nuclear localization signal (24, 25), which led to an accumulation of fluorescence in the nucleus of injected cells within 10 – 20 minutes (Fig. 1b). As shown previously, ProTα does not exhibit significant attractive interaction with the intracellular environment (13), making it an ideal candidate for probing generic crowding effects in cells. Compared to *in vitro* conditions, the dimensions of ProTα were unchanged upon injection into the cells, but translational diffusion was slower by a factor of ~2 in the cytosol and nucleus, as expected from an increased intracellular viscosity. These observations suggest that crowding effects under the usual near-physiological growth conditions of HeLa cells are too small to induce detectable changes in ProTα dimensions. In previous crowding experiments in vitro, however, ProTα exhibited significant compaction even in the presence of a few percent volume fraction of PEG. Even moderate changes in intracellular concentrations might thus be sufficient for reaching a regime where crowding effects on ProTα become detectable.

To increase intracellular crowding, HeLa cells were thus subjected to an osmotic shock using PEG 400, an uncharged hydrophilic polymer that cannot cross the membrane (26). Upon addition of 20 % PEG 400 in the extracellular medium, a collapse of the adherent cells (Fig. 1c) by approximately 50 % was observed within less than 2 min, and the contraction was stable for at least 30 min (Fig. 1d). Injection of ProTα into the osmotically shocked cells showed that the altered intracellular environment influenced the dimensions and dynamics of the IDP, reflected by an increase in transfer efficiency of ProTα. The mean transfer efficiency of ProTα increased from 0.34 ± 0.01 in the cytosol of cells under physiological conditions to 0.39 ± 0.02 in the cytosol of crowded cells (Fig. 2a), corresponding to a change in radius of gyration (R_g) of ~10 % (Fig. 2d). The strongest effect of the crowded intracellular environment was observed on the translational diffusion of ProTα. The diffusion time (τ_d) of ProTα in the crowded cytosol was 4.6-times slower (6.9 ms ± 1.9 ms) than in the cytosol at physiological conditions (1.5 ms ± 0.2 ms) (Fig. 2b and 2d). Surprisingly, the physiological and crowded cytosol had a comparable effect on the chain dynamics of ProTα (Fig. 2c). In both cases, the chain reconfiguration times are approximately 2.3-times slower than *in vitro* (80 ns ± 11 ns at physiological conditions, 82 ns ± 13 ns in the crowded cytosol) (Fig. 2d).

These results lead to two conclusions: First, the concentration of macromolecules in the cytosol of HeLa cells under physiological conditions is so low that it impedes diffusion of ProTα only very little and has no detectable effect on the dimensions of the IDP. We thus analyzed the protein concentrations in several cell lines (Tab. *1*) and found total protein concentrations between 64 and 123 mg/ml, within the range previously reported for eukaryotic cells (2, 15). As revealed by experiments where only cytoplasmic proteins were quantified, the concentration of soluble cytosolic proteins, however, was much lower (25 mg/ml in HEK 293-F cells). This result was not unexpected, since a lot of the intracellular proteins are membrane-associated (27-29). Second, we concluded that the influence of the intracellular environment on the dynamics and dimensions of proteins is highly length-scale dependent. We found a small effect of the cytosol on the chain dynamics and dimensions of ProTα both for physiological and crowded cells (Fig. 2d), properties for which the relevant length-scales are in the sub-nanometer (chain dynamics) or low nanometer regime (dimensions). The largest effect was observed for translational diffusion, where the relevant length-scale is in the micrometer regime, determined by the size of the observation volume of the confocal microscope.

To further investigate the length-scale dependence, we injected two other molecules with very different sizes into HeLa cells: the fluorophore Atto 532 (Mw = 900 Da, R_h = ~0.5 nm) and the homotetrameric protein β-glucuronidase (GUS; Mw = 280 kDa, R_h = 5.1 nm) labeled with Atto 532 (Fig. 3a). ProTα (R_h = 4.4 nm) and GUS have a similar R_h in spite of their large difference in molecular mass because GUS is a globular folded protein, whereas ProTα is completely unstructured and thus expanded (30). Upon injection into HeLa cells, Atto 532 exhibited a similar relative slowdown in translational diffusion as ProTα: In the cytosol under physiological conditions, diffusion was 1.7-times slower than *in vitro* (ProTα: 1.6-times slower) (Fig. 3b), whereas in crowded cells, the diffusion was 7.6-times slower (ProTα: 7.3-times slower). In contrast, GUS diffuses 2.8 times more slowly in the cytosol and 15 times more slowly in the crowded cytosol. The similarity in translational diffusion of Atto 532 and ProTα combined with the dissimilarity of GUS and ProTα suggests that the IDP can move through a crowded environment much more efficiently than the rigid protein GUS, most likely by reptation through the meshwork of entangled macromolecules.

To assess the influence of the size of crowders on the diffusional behavior of proteins, we quantified translational diffusion of ProTα *in vitro* in the presence of different PEGs and Ficoll (Fig. 4a). Ethylene glycol (EG) and short PEGs, such as PEG 400, resulted in only little retardation in solutions of up to 20 % (w/v) crowder. Larger PEGs reduced diffusion approximately 6-fold with PEG 6000 and 9-fold with PEG 35'000 at concentrations of ~20 % (w/v). The diffusivity of ProTα in PEG solutions was analyzed by assuming a depletion layer around the protein, which determines how fast translational diffusion is in relation to the size and viscosity of the crowder (31) (solid lines, Fig. 4a). This approach described the crowder concentration dependence of the diffusion of ProTα in solutions with the linear polymer PEG and branched Ficoll 70/400 relatively accurately. However, the apparent radius of ProTα had to be adjusted to a value smaller than the hydrodynamic radius measured in dilute solution (R_h = 1.8 nm instead of 4.4 nm), although the radius of gyration estimated from intramolecular FRET experiments on ProTα is only reduced by ~10 % upon crowding under these conditions (23). This finding supports the idea that highly flexible IDPs such as ProTα can easily diffuse through a crowded solution, resulting in a reduced apparent hydrodynamic radius.

In spite of the obvious heterogeneity of the environment *in vivo*, it would be of interest whether conditions can be reconstituted *in vitro* that approximate the cellular situation in terms of diffusion and dynamics across length- and timescales. Interestingly, if the diffusion of ProTα, Atto 532, and GUS is measured in PEG 35'000 as a crowder at concentrations of about 5 % (w/v) (Fig. 4b), we observe very similar behavior to what we found in the cytosol of cells, with a similar slowdown in translational diffusion for the three tested molecules as in the cytosol of HeLa cells. Intracellular conditions of crowded cells can be mimicked by approximately 20 % (w/v) PEG 35'000, yielding a comparable diffusional retardation. Additionally, the chain reconfiguration time of ProTα in 20 % (w/v) PEG 35'000 (97 ns ± 5 ns) was similar to the reconfiguration time in the cytosol of crowded cells (82 ns ± 13 ns).

Taking together, we conclude that the cytosol of commonly used eukaryotic cell lines is less crowded than commonly assumed, allowing proteins and IDPs to easily diffuse in the cell. However, increasing the intracellular concentration of macromolecules by only a factor of two by osmotic stress increases crowding to an extent that leads to substantially slowed diffusion and detectable compaction of IDPs such as ProTα. Solutions of 5 to 20 % (w/v) PEG 35’000 provide a reasonable in-vitro mimic of the diffusional dynamics of proteins and small molecules and may thus resemble the average crowder size and concentration of the eukaryotic cytosol. Interestingly, current estimates of the average size of eukaryotic proteins are in the range of 350 - 450 amino acids (32), corresponding to a hydrodynamic radius which is similar to the size of PEG 35’000.

Methods

Expression and labeling of prothymosin α (ProTα)

ProTα was expressed recombinantly in *Escherichia coli* BL21 and purified with a Ni-NTA column as previously described (13). Cysteine residues for the specific labeling of ProTα using maleimide chemistry were introduced by site-directed mutagenesis at positions 1 and 56. ProTα was sequentially labeled with Atto 532 maleimide (ATTO-TEC GmbH) as a donor and Biotium CF680R maleimide (Biotium Inc.) as an acceptor dye. The Förster radius for the dye pair was calculated to be 5.7 nm.

Single-molecule FRET measurements in HeLa cells and *in vitro*

Instrumentation

Single-molecule fluorescence measurements were performed on a Microtime 200 instrument (PicoQuant) using an Olympus IX71 microscope body (13). The sample was excited using a 532-nm continuous-wave laser (LaserBoxx LMX-532S, Oxxius) at 50 μW of excitation power, as measured at the back aperture of the objective. The laser was focused into the sample with an Olympus UplanApo 60×/1.20-W objective (Olympus), and emitted photons from the sample were collected through the same objective. Scattered excitation light was eliminated with a long-pass filter (LP532, Chroma Technology) before the emitted photons passed a 100-μm confocal pinhole. The emitted photons were then separated into four channels with a polarizing beam splitter and two dichroic mirrors (635DCXR, Chroma Technology). Donor photons were filtered using an ET585/65m filter (Chroma Technology), acceptor photons were filtered using LP647RU and HC750/SP filters (Chroma Technology). Photons from each channel were detected using four SPCM-AQR-15 single-photon avalanche diodes (PerkinElmer). The arrival times of detected photons were recorded with four channels of a HydraHarp 400 counting module (PicoQuant) with a resolution of 16 ps. The objective was mounted on a piezo stage combination (P-733.2 and PIFOC, Physik Instrumente GmbH) to enable 3D scans.

Cell culture

HeLa cells (American Type Culture Collection) were cultured as previously described (13). One day before a single-molecule experiment, 2×10^5 HeLa cells in 4 ml phenol red–free DMEM medium were pipetted into a microscopy dish (35-mm glass-bottom dishes, Ibidi; prior to the experiment, the glass-surface was coated with poly(l-lysine) solution (0.01% in H2O, Sigma)) and incubated overnight (37 °C, 5% CO_2) until the cells became adherent again. Prior to single-molecule measurements, the medium was exchanged with Hanks balanced salt solution (HBSS without Ca^{2+}/Mg^{2+}, Sigma).

Microinjection.

Protein solutions were injected into HeLa cells using a Xenoworks microinjection system (Sutter Instrument) (13). The pressure pulses (0.3 - 0.5 bar, 30 - 50 ms) for the injections were delivered by a home-built injector. The sample was injected into cells using Femtotips II capillaries (Eppendorf). All protein samples for injections contained 0.005% Tween 20 (Pierce) to reduce adsorption of protein to the capillary and to any container surfaces used during dilution, and the sample was injected into a targeted cell. Solutions for injection contained labeled protein at concentrations of 10 - 20 nM.

In-cell FRET and nanosecond fluorescence correlation spectroscopy (nsFCS) of ProTα

Single-molecule FRET efficiency histograms of ProTα in living cells were acquired by injecting samples with a protein concentration of 10 nM into the cytosol or nucleus of HeLa cells (13). All experiments with ProTα were performed at 22 °C, and each data point was collected in a different cell. For experiments under crowded cellular conditions after osmotic shock, the cells were incubated with 20 % PEG 400 in HBSS just before the experiment. For subsequent experiments, the crowded cells were exchanged with fresh osmotically shocked cells every 45 minutes. Successful injections were verified by recording fluorescent images of the cells in axial and horizontal direction before and after the injections. The confocal volume was positioned either in the cytosol or the nucleus of an injected cell and away from regions of high cellular autofluorescence. For each injected cell, data were acquired for 3 min, until most of the injected protein was photobleached.

nsFCS measurements in cells were performed by injecting 20 nM ProTα in HBSS (without Ca^{2+}/Mg^{2+}), containing 0.005% Tween 20. Injections and measurements were performed as described above. Owing to the poorer photon statistics for interphoton times in the nanosecond range, the measurements require a substantially larger total number of photons to be detected than are needed for FRET efficiency histograms. To achieve sufficient statistics, we averaged the results of 51 single-cell measurements under physiological conditions (total acquisition time of 2.5 h, data taken from König *et al.* 2015 (13)) and of 40 single-cell measurements under crowded/osmotically shocked conditions (total acquisition time of 2.0 h). Control measurements in buffer were carried out with protein concentrations of 100 pM in HBSS (without Ca^{2+}/Mg^{2+}) with 0.001% Tween 20 for 2 h.

FCS measurements to determine the translational diffusion of ProTα, Atto 532 and β-glucuronidase in different PEG and Ficoll solutions (ethylene glycol, PEG 400, PEG 6000, PEG 35'000; Ficoll 70, Ficoll 400) were carried out with sample concentrations of 100 pM in HBSS (without Ca^{2+}/Mg^{2+}) with 0.001% Tween 20 and PEG or Ficoll concentrations between 0 – 20 % (w/v) at pH 7.2. Data were recorded for 30 min for each sample.

Data analysis

Transfer efficiency histograms

Fluorescence bursts from individual molecules were identified by binning the photon trajectory into 1-ms bins. Bins with a minimal threshold of 40 – 100 photons, chosen depending on the intensity of the background signal, and a maximum of 10'000 photons were considered as a burst. Contiguous bins fulfilling these criteria were combined into one burst. Since the background was variable during an in-cell measurement owing to photobleaching, the photon trajectory was split into shorter intervals where the background was approximately constant. Detected bursts from each interval were summed up to construct the final transfer efficiency histogram. The transfer efficiency of each burst was calculated according to $E = n_A/(n_A + n_D)$, where n_D and n_A are the numbers of donor and acceptor photons, respectively, corrected for background in the corresponding interval, acceptor direct excitation, channel cross-talk, differences in detector efficiencies, and quantum yields of the dyes as previously described (13).

As described recently (13), subpopulations in transfer efficiency histograms were approximated with Gaussian (G) and four-parameter log-normal (L) peak functions for symmetric and asymmetric peaks. Transfer efficiency histograms of ProTα were fitted using a log-normal peak function for the donor-only population and a Gaussian peak function for the FRET-labeled population without constraints of fit parameters.

Interdye distances and radii of gyration

Dimensions of ProTα were calculated from the measured transfer efficiencies as described earlier (13). Briefly, the mean transfer efficiency, $\langle E\rangle$, can be expressed as a function of the distance dependence of the transfer efficiency, $E(r)$, weighted for the dye-to-dye distance distribution, $P(r)$, sampled by the chain,

$$\langle E\rangle = \int_a^{l_c} E(r)P(r)dr$$

with

$$E(r) = \frac{1}{1+(r/R_0)^6}$$

where a is the distance of closest approach of the two dyes, l_c is the contour length of the labeled protein segment, and R_0 is the Förster radius of the chosen dye pair. We assumed a Gaussian chain model for the distribution of distances, $P(r)$,

$$P(r) = 4\pi r^2\left(\frac{3}{2\pi\langle r^2\rangle}\right)^{3/2} exp\left(-\frac{3r^2}{2\langle r^2\rangle}\right)$$

where $\langle r^2\rangle$ is the mean squared interdye distance (33). These equations were solved for $\langle r^2\rangle$ numerically and then converted to the mean radius of gyration, R_g, of the chain using

$$R_g^2 = \frac{\langle r^2\rangle}{6}.$$

Fluorescence correlation spectroscopy (FCS)

The diffusion time of the labeled proteins was extracted from FRET experiments by correlating the fluorescence intensity fluctuations using

$$g_{ij}(\tau) = \frac{\langle n_i(t)n_j(t+\tau)\rangle_t}{\langle n_i(t)\rangle_t\langle n_j(t)\rangle_t}$$

with *i, j* = A, D (A = acceptor, D = donor), and where $n_i(t)$ and $n_j(t+\tau)$ are the fluorescence count rates in the detector channels i and j at time t and after a lag time τ, respectively. $\langle \ldots \rangle_t$ denotes the time average of the detected signal. The resulting cross-correlation curves were fitted according to (13) to obtain the translational diffusion time τ_D.

nsFCS

Donor and acceptor autocorrelation curves and cross-correlation curves between acceptor and donor channels were calculated as described earlier (13, 34, 35). The resulting auto- and cross-correlation curves were fitted globally up to a delay time of

1 µs, and the extracted time constant of the chain dynamics was converted into a chain reconfiguration time τ_r according to (36).

Analysis of the diffusion of proteins in crowded solutions

The diffusion of proteins in crowded solutions was described using the approach of Tuinier *et al.* 2006 (31) and Lekkerkerker and Tuinier 2011 (37). According to Stokes' law, the diffusion coefficient D of an object is

$$D = \frac{kT}{f},$$

with the friction coefficient $f = 6\pi\eta a$, where *a* is the equivalent spherical radius, and η is the viscosity of the medium. However, in a crowded solution, an object experiences a different friction than expected from the bulk viscosity of the solution if we account for the depletion of macromolecules surrounding it. In this case, the effective friction coefficient becomes

$$f_t = 6\pi\eta_s a \frac{Q(\lambda,\epsilon)}{Z(\lambda,\epsilon)}, \text{ with } \lambda = \frac{\eta_s}{\eta_m}, \epsilon = \frac{\delta}{a}.$$

η_s corresponds to the solvent viscosity (1 *cP* for water), and η_m is the bulk viscosity of the solution. δ corresponds to the thickness of the depletion layer. The numerator and denominator of the fraction in the effective friction coefficient is calculated as follows:

$$\begin{aligned} Q(\lambda,\epsilon) &= 2(2+3\lambda)(1+\epsilon)^6 - 4(1-\lambda)(1+\epsilon), \\ Z(\lambda,\epsilon) &= 2(2+3\lambda)(1+\epsilon)^6 - 9\left(1 - \frac{\lambda}{3} - \frac{2\lambda^2}{3}\right)(1+\epsilon)^5 + 10(1-\lambda)(1+\epsilon)^3 \\ &\quad - 9(1-\lambda)(1+\epsilon) + 4(1-\lambda)^2. \end{aligned}$$

The depletion layer thickness δ for dilute and semi-dilute polymer solutions at a hard wall is

$$\delta^{-2} = \delta_0^{-2} + \xi^{-2},$$

with ξ being the correlation length ($\xi = Rg\left(\frac{c}{c^*}\right)^{-3/4}$, $c*$ is the overlap concentration), and δ_0 is the depletion thickness for dilute polymer solutions ($\delta_0 = 1.07R_g$). The depletion layer thickness at a hard wall is then converted into the depletion layer thickness around a sphere, δ_s.

$$\delta_s = \left([1 + c_1 q + c_2 q^2 + c_3 q^3]^{1/3} - 1\right) \cdot a,$$

with $q = \frac{\delta}{a}$, $c_1 = 3$, $c_2 = 2.273$, and $c_3 = -0.0975$ (37). The depletion layer thickness around a sphere was then used to calculate the friction term f_t.

Determination of cellular protein and DNA concentrations

The total protein concentration of different mammalian cell lines was determined by cultivating HeLa, HEK 293 wt, HEK 293 F, COS 7, and CHO K1 (American Type Culture Collection) in 150 ml tissue culture flasks in their optimal medium (DMEM for HeLa, HEK 293 wt/F and COS-7 cells; Ham's F12 medium for CHO-K1 cells), until approximately 0.5 ml of wet cell pellet of each cell line were obtained. The cells were lysed, and the protein concentration was determined as described by (38). Briefly, the cell pellets were lysed in a 0.1 M Tris buffer pH 7.8, containing 2% SDS and 50 mM DTT, by vortexing them intensely and heating the samples to 90 °C. The sample was then diluted into 0.1 M Tris pH 8.5, 8 M urea, and the tryptophan fluorescence was measured (excitation at 295 nm, emission detected at 350 nm). Based on calibration solutions of tryptophan, the tryptophan concentration in the sample was calculated. Using an average tryptophan content in eukaryotic cell tissue of 1.17 % (by weight) (38), the total protein concentration of the cell pellet was estimated. The total DNA concentration was estimated by measuring the absorbance at 260 nm and using an extinction coefficient of 0.02 $(\mu g/ml)^{-1}cm^{-1}$ (39).

For the determination of the cytosolic protein concentration, HEK 293 F cells were cultivated in suspension culture to obtain a higher number of cells. The HEK 293 F cells were cultivated until 5 ml of wet cell pellet was obtained. The cell pellet was processed according to Abmayr et al. (40). The protein concentration in the obtained cytoplasmic extract was analyzed using a BCA protein assay.

Figures

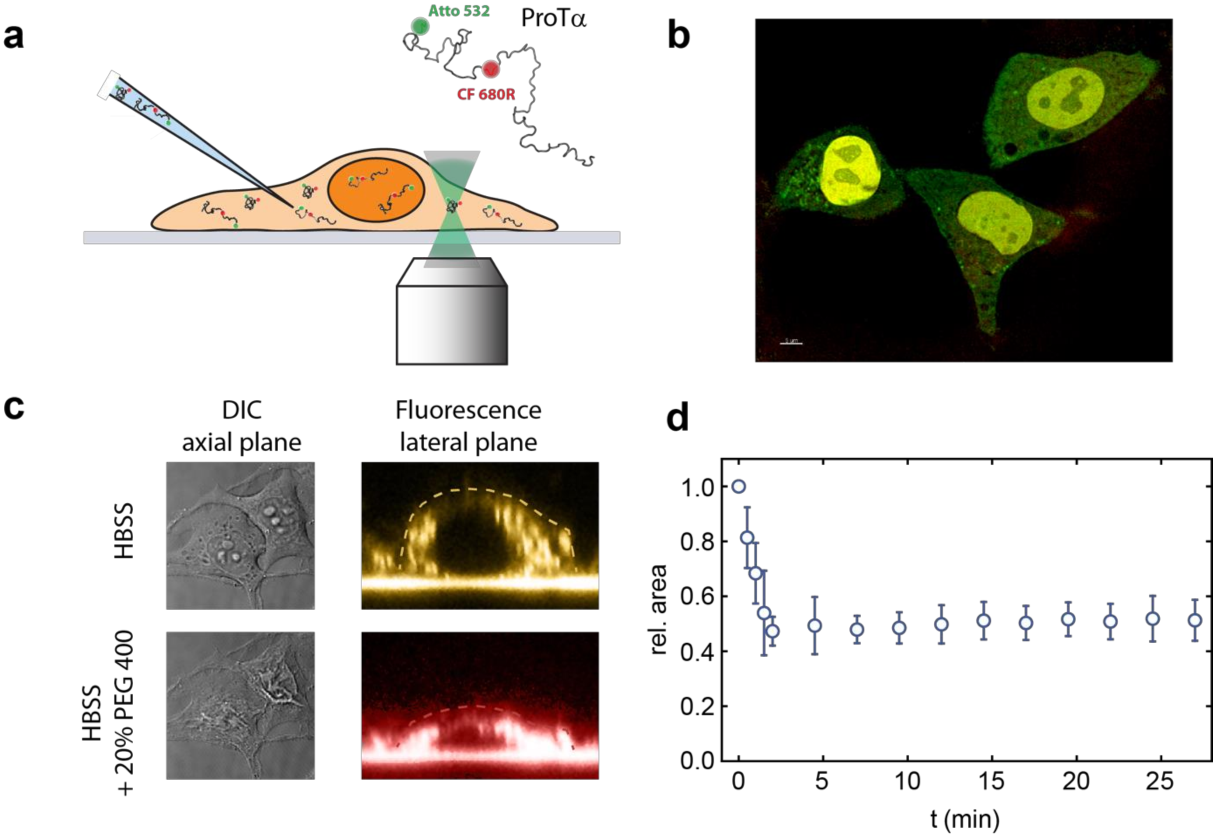


Fig. 1: Single-molecule fluorescence measurements in live eukaryotic cells under physiological and crowded conditions.

a) Schematic representation of the microinjection of fluorescently labeled Prothymosin α (ProTα) into adherent HeLa cells.

b) Confocal fluorescence image of HeLa cells, showing the distribution of labeled ProTα after microinjection of the protein into the cytosol. Due to a C-terminal nuclear localization sequence, ProTα becomes enriched in the nucleus within ~10 min.

c) Differential interference contrast (DIC) and confocal fluorescence images of HeLa cells before and after osmotic shock with PEG 400, showing the compaction of cells in the axial (DIC) and lateral planes (fluorescence). The contour of the cells in the fluorescence images is indicated with dashed lines.

d) Plot of the relative area of the cell in the vertical plane after the osmotic shock (8 individual cells, mean value and s.d. plotted). The relative area reduced to approximately half of its initial value and remained constant for at least 30 min.

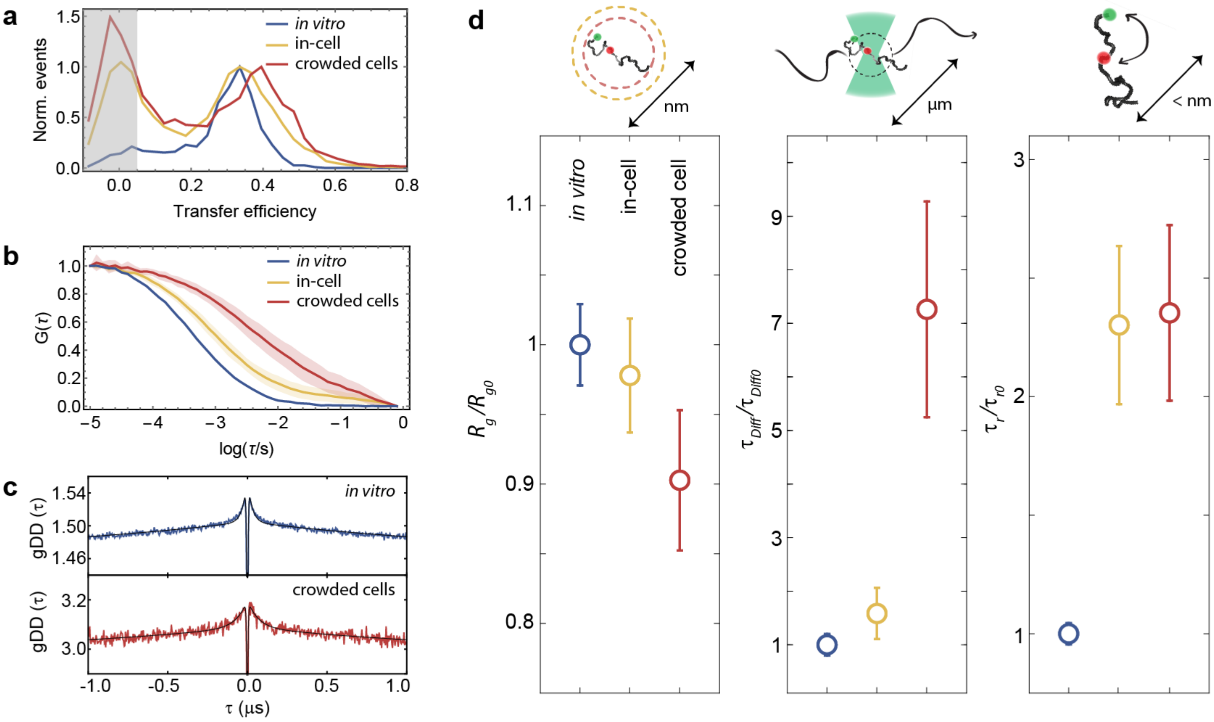


Fig. 2: Effect of intracellular crowding on the dimensions and dynamics of ProTα.
a) Transfer efficiency histograms on ProTα *in vitro* (blue; 0.33 ± 0.01), in the cytosol of HeLa cells (yellow; sum of 8 individual histograms, 0.34 ± 0.01), and in the cytosol of crowded HeLa cells (red; sum of 6 individual histograms, 0.39 ± 0.02). Measurements of the transfer efficiency in the nucleus of cells under physiological (0.35 ± 0.03, n= 8) and crowded cells (0.40 ± 0.02, n = 5) yielded similar results as in the corresponding cytosol.
b) FCS cross-correlation curves of ProTα *in vitro* (blue; 0.9 ms ± 0.2 ms), in the cytosol of HeLa cells (yellow; 1.5 ms ± 0.2 ms, n = 8), and in the cytosol of crowded HeLa cells (red; 6.9 ms ± 1.9 ms, n = 6) show the slowdown in translational diffusion with increasing intracellular crowding. The shaded bands represent the standard deviations of the measurements. The corresponding FCS measurements in nuclei yielded similar results (HeLa cells: 1.8 ms ± 0.8 ms, n = 8; crowded cells: 6.5 ms ± 2.5 ms, n = 7).
c) nsFCS curves (donor autocorrelation) of ProTα *in vitro* (blue) and in the cytosol of crowded cells (red) show that the chain reconfiguration dynamics experience a slowdown in HeLa and in crowded cells. The chain reconfiguration time in vitro was 35 ns ± 1 ns, in the crowded cytosol 82 ns ± 13 ns.
d) Plot of the relative dimensions (R_g/R_{g0}), relative diffusion times (τ_{Diff}/τ_{Diff0}), and relative chain reconfiguration times (τ_r/τ_{r0}) of ProTα *in vitro* (blue), in HeLa cells (yellow) and in crowded cells (red). The icon on top of each panel indicates the length-scales on which the chain dimensions (nm), translational diffusion (µm), and chain dynamics (< nm) are affected by other (macro-)molecules.

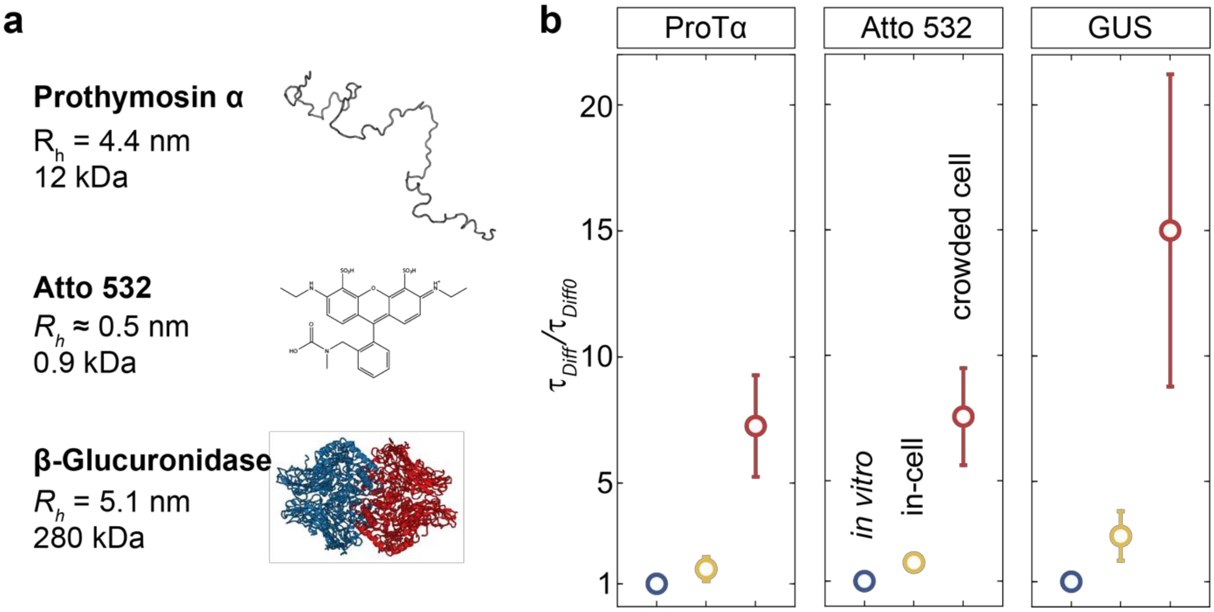


Fig. 3: Translational diffusivity of molecules of different sizes in the cytosol of cells.
a) Hydrodynamic radii and molecular masses of ProTα, Atto 532, and β-glucuronidase (GUS) with their structures.
b) Plot of the relative diffusion times of ProTα, Atto 532, and GUS *in vitro* (blue), in-cell (yellow) and in crowded cells (red).

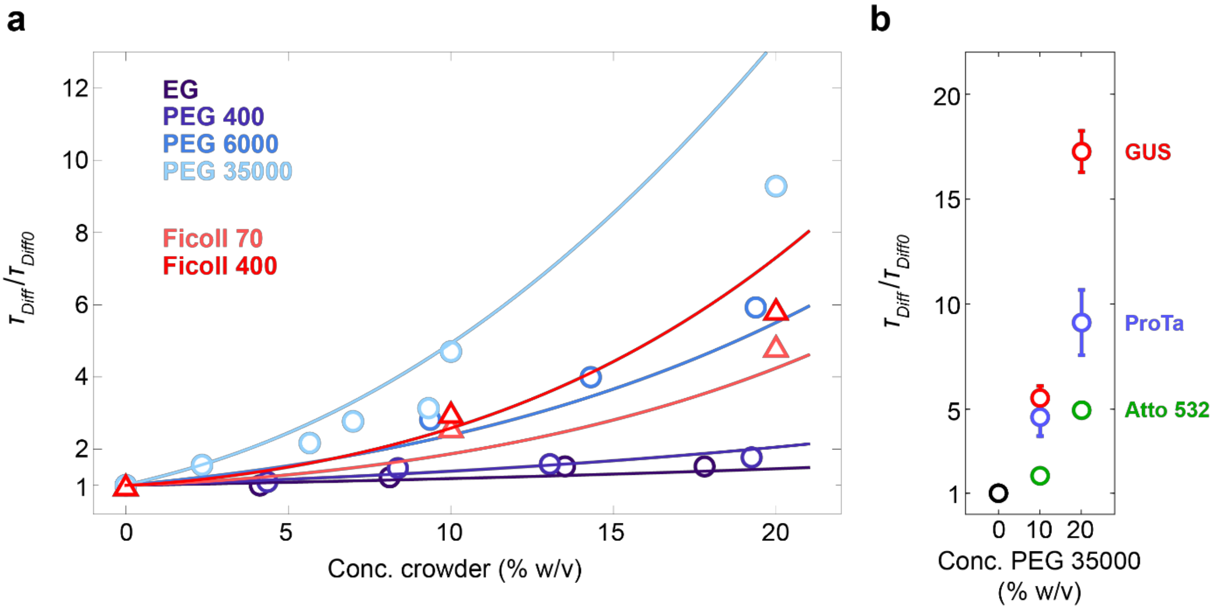


Fig. 4: Translational diffusion of proteins in crowded solutions in vitro.
a) Relative translational diffusion of ProTα in solutions of EG (dark purple circles), PEG 400 (purple circles), PEG 6000 (blue circles), PEG 35’000 (light blue circles). Additional experiments were done in solutions with the branched polymer Ficoll 70 (light red triangles) and Ficoll 400 (red triangles). The solid lines represent the calculated diffusion time (31) of ProTα in each crowder solution, assuming an apparent R_g of ProTα of 1.8 nm.
b) The relative diffusion times of Atto 532 (green), ProTα (blue), and GUS (red) in 10 and 20 % (w/v) PEG 35’000.

Tab. 1: Total protein and nucleic acid concentrations of four commonly used adherent cell lines (HeLa, HEK 293 wt, COS-7, CHO-K1) and one suspension cell line (HEK 293-F). From the latter cell line, also the cytosolic protein concentration was determined.

Cell line	Total protein		Cytosolic protein		Nucleic acids	
	Mean (mg/ml)	S.D. (mg/ml)	Mean (mg/ml)	S.D. (mg/ml)	Mean (mg/ml)	S.D. (mg/ml)
HeLa	108	*25*	-	-	23	*5*
HEK 293 wt	90	32	-	-	11	*2*
HEK 293-F	108	28	25	5	18	*4*
COS-7	64	15	-	-	21	*4*
CHO-K1	123	12	-	-	25	*5*

References

1. Zimmerman SB, Trach SO (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. *J Mol Biol* 222(3):599–620.
2. Cheung MC, et al. (2013) Intracellular protein and nucleic acid measured in eight cell types using deep-ultraviolet mass mapping. *Cytom Part A* 83(6):540–551.
3. Ittah V, Kahana E, Amir D, Haas E (2004) Applications of time-resolved resonance energy transfer measurements in studies of the molecular crowding effect. *J Mol Recognit* 17(5):448–455.
4. Dhar A, et al. (2010) Structure, function, and folding of phosphoglycerate kinase are strongly perturbed by macromolecular crowding. *Proc Natl Acad Sci USA* 107(41):17586–17591.
5. Wang Y, Benton LA, Singh V, Pielak GJ (2012) Disordered Protein Diffusion under Crowded Conditions. *J Phys Chem Lett* 3(18):2703–2706.
6. Mittal S, Singh LR (2013) Denatured state structural property determines protein stabilization by macromolecular crowding: a thermodynamic and structural approach. *PLOS ONE* 8(11):e78936.
7. Kuhn T, et al. (2011) Protein Diffusion in Mammalian Cell Cytoplasm. *PLOS ONE* 6(8):e22962.
8. Dhar A, et al. (2011) Protein stability and folding kinetics in the nucleus and endoplasmic reticulum of eucaryotic cells. *Biophys J* 101(2):421–430.
9. Theillet F-X, et al. (2016) Structural disorder of monomeric alpha-synuclein persists in mammalian cells. *Nature* 530:45–50.
10. Crawford R, et al. (2013) Long-lived intracellular single-molecule fluorescence using electroporated molecules. *Biophys J* 105(11):2439–2450.
11. Murakoshi H, et al. (2004) Single-molecule imaging analysis of Ras activation in living cells. *Proc Natl Acad Sci USA* 101(19):7317–7322.
12. Sakon JJ, Weninger KR (2010) Detecting the conformation of individual proteins in live cells. *Nat Methods* 7(3):203–205.
13. König I, et al. (2015) Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells. *Nat Methods* 12(8):773–779.
14. McGuffee SR, Elcock AH (2010) Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLOS Comput Biol* 6(3):e1000694.
15. Conlon I, Raff M (2003) Differences in the way a mammalian cell and yeast cells coordinate cell growth and cell-cycle progression. *J Biol* 2(1). doi:10.1186/1475-4924-2-7.
16. Rivas G, Minton AP (2016) Macromolecular Crowding In Vitro, In Vivo, and In Between. *Trends in Biochemical Sciences* 41(11):970–981.
17. Boersma AJ, Zuhorn IS, Poolman B (2015) A sensor for quantification of macromolecular crowding in living cells. *Nat Methods* 12(3):227–229.

18. Gnutt D, Gao M, Brylski O, Heyden M, Ebbinghaus S (2015) Excluded-volume effects in living cells. *Angew Chem Int Ed* 54(8):2548–2551.

19. Miklos AC, Sarkar M, Wang Y, Pielak GJ (2011) Protein crowding tunes protein stability. *J Am Chem Soc* 133(18):7116–7120.

20. Sarkar M, Smith AE, Pielak GJ (2013) Impact of reconstituted cytosol on protein stability. *Proc Natl Acad Sci USA* 110(48):19342–19347.

21. Sarkar M, Lu J, Pielak GJ (2014) Protein crowder charge and protein stability. *Biochemistry* 53(10):1601–1606.

22. Brucale M, Schuler B, Samorì B (2014) Single-molecule studies of intrinsically disordered proteins. *Chem Rev* 114(6):3281–3317.

23. Soranno A, et al. (2014) Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments. *Proc Natl Acad Sci USA* 111(13):4874–4879.

24. Gomez-Marquez J, Segade F (1988) Prothymosin α is a nuclear protein. *FEBS Lett* 226(2):217–219.

25. Manrow RE, Sburlati AR, Hanover JA, Berger SL (1991) Nuclear targeting of prothymosin alpha. *Biol Chem* 266(6):3916–3924.

26. Ma TY, Hollander D, Krugliak P, Katz K (1990) PEG 400, a hydrophilic molecular probe for measuring intestinal permeability. *Gastroenterology* 98(1):39–46.

27. Lander ES, et al. (2001) Initial sequencing and analysis of the human genome. *Nature* 409(6822):860–921.

28. Ahram M, Litou ZI, Fang R, Al-Tawallbeh G (2006) Estimation of membrane proteins in the human proteome. *In Silico Biol* 6(5):379–386.

29. Spitzer J, Poolman B (2013) How crowded is the prokaryotic cytoplasm? *FEBS Lett* 587(14):2094–2098.

30. Müller-Späth S, et al. (2010) Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc Natl Acad Sci USA* 107(33):14609–14614.

31. Tuinier R, Dhont JKG, Fan TH (2006) How depletion affects sphere motion through solutions containing macromolecules. *Europhys Lett* 75(6):929–935.

32. Brocchieri L, Karlin S (2005) Protein length in eukaryotic and prokaryotic proteomes. *Nucleic Acids Res* 33(10):3390–3400.

33. Schuler B, Soranno A, Hofmann H, Nettels D (2016) Single-Molecule FRET Spectroscopy and the Polymer Physics of Unfolded and Intrinsically Disordered Proteins. *Annu Rev Biophys* 45(1):207–231.

34. Hillger F, et al. (2008) Probing protein-chaperone interactions with single-molecule fluorescence spectroscopy. *Angew Chem Int Ed* 47(33):6184–6188.

35. Nettels D, Gopich IV, Hoffmann A, Schuler B (2007) Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc Natl Acad Sci USA* 104(8):2655–2660.

36. Gopich IV, Nettels D, Schuler B, Szabo A (2009) Protein dynamics from single-molecule fluorescence intensity correlation functions. *J Chem Phys*

131(9):095102.

37. Lekkerkerker HNW, Tuinier R (2011) *Colloids and the Depletion Interaction* (Springer Science & Business Media) doi:10.1007/978-94-007-1223-2.
38. Wiśniewski JR, Gaugaz FZ (2015) Fast and sensitive total protein and Peptide assays for proteomic analysis. *Anal Chem* 87(8):4110–4116.
39. Surzycki S (2012) *Basic Techniques in Molecular Biology* (Springer Science & Business Media).
40. Abmayr SM, Yao T, Parmely T, Workman JL (2006) Preparation of nuclear and cytoplasmic extracts from mammalian cells. *Curr Protoc Mol Biol* Chapter 12:Unit 12.1–12.1.10.

6 Conclusions and Outlook

Most experiments to elucidate the structures and functions of proteins have been carried out in the test tube. This approach has provided enormous insight into the nature of proteins. However, studying proteins *in vitro* does not represent their natural environment, which is in the highly complex and crowded environment of living cells. A particularly interesting class of proteins in this context are IDPs because of their partial or complete lack of secondary and tertiary structure, which is expected to make their conformations and dynamics particularly susceptible to crowding and other cellular effects. A dogma of structural biology was that 3D-structure determines protein function. In the case of IDPs, this idea is no longer completely valid. Despite their disordered nature, IDPs still interact with other proteins or ligands in a specific way. Especially because of their highly dynamic nature, they are, e.g., part of large signaling networks to regulate transcription (1). Studying the effect of the crowded cellular environment on the structures and dynamics of IDPs may thus provide important insights into protein function in the cell. Here, we approached this question in two steps: first, by adding polymers of defined size and concentration to mimic cellular crowding, and second, directly in live cells.

For the first step, single-molecule FRET spectroscopy was employed to study the effects of different co-solutes on the structures and dynamics of IDPs. In an initial study, the influence of macromolecular crowding on the conformations of the disordered proteins ProTα, ACTR and the N-terminal domain of HIV-1 integrase (IN) was analyzed (2). Using artificial crowders such as PEG, we found that all IDPs adopted increasingly compact conformational distributions with increasing concentration and size of the PEG polymers. Scaled particle theory (SPT), the standard theory for crowding, fails to describe the observed effects because it assumes the IDPs and the crowders to be hard spheres. Rather, we found that both the IDPs and the crowders need to be treated as polymeric chains which can interpenetrate to explain our observations. By adapting the Flory-Huggins theory to the short and the long chain regime, it was possible to quantify the effects of crowders with different sizes and concentrations on IDPs.

The presence of high concentrations of crowders modulates the dimensions of IDPs. Since many molecules in a living cell are polymers, such as nucleic acids, polysaccharides or other disordered proteins with estimated intracellular concentrations of up to 300 mg/ml (3, 4), they might have a similar crowding effect on IDPs. Additionally, the compartmentalization of a cell can lead to very high local concentrations of macromolecules, such as along cell membranes (5) or in RNA granules and other intracellular protein bodies (6).

Studying proteins in living eukaryotic cells has remained challenging due to the high complexity of the intracellular environment. Amongst others, fluorescence spectroscopy methods have been shown to be very useful for studying intracellular molecules because of their specificity for fluorescent labels. The application of ensemble FRET has already provided insights into, e.g., protein folding stabilities and dynamics (7) and the effect of intracellular crowding on protein sensors (8). First FRET experiments with single-molecule resolution at the membrane and within eukaryotic cells were reported by Murakoshi *et al.* (9) and by Sakon & Weninger (10), respectively, and the group of Kapanidis successfully established single-molecule FRET measurements in bacterial cells (11). Using electroporation for the delivery of labeled molecules, single-molecule FRET in living bacteria has allowed the study of DNA replication, transcription and repair.

To unfold the full potential of single-molecule FRET spectroscopy in living eukaryotic cells, several requirements had to be fulfilled. Most previous in-cell FRET experiments relied on fluorescent proteins. Because of their relatively large size and the low brightness and photostability, small organic dyes with superior photophysical properties are preferred. Furthermore, ways to reduce cellular autofluorescence and to deliver sub-micromolar concentrations of sample into cells had to be found. These problems were approached by Sakon & Weninger in 2010 (10), where they used microinjection to deliver fluorescently labeled proteins into single cells and detected them with total internal reflection fluorescence (TIRF) microscopy. We further optimized this approach by combining confocal single-molecule FRET spectroscopy with microinjection to deliver the sample into adherent HeLa cells (12). Microinjection turned out to be a good method for reproducibly loading cells with picomolar to nanomolar concentrations of protein. It allowed us to inject proteins specifically into the cytosol or the nucleus of a cell, while maintaining a high survival rate of injected cells. Also, the use of red-shifted fluorescent dyes as FRET reporters, such as Atto 532 / Biotium CF680R or Atto 532 / Abberior Star 635, allowed us to excite the injected molecules with a wavelength of 520 nm to minimize the prohibitive influence of intracellular auto-fluorescence. The data obtained from single cells were of good statistical quality, which enabled the analysis of single-cell FRET histograms. From the same data, we also obtained information about the translational diffusion of proteins in cells by FCS. By increasing the concentration of injected sample, it was even possible do nsFCS measurements in the cell, which allowed us to study protein chain dynamics in the submicrosecond regime.

We injected the intrinsically disordered protein ProTα into adherent HeLa cells and showed that we can measure the dimensions of IDPs in the cytosol and the nucleus of living cells. ProTα showed the same highly expanded dimensions in the cytosol and the nucleus of cells as under *in vitro* conditions at physiological ionic strength, suggesting that ProTα remains unstructured within cells. This finding was further supported by nsFCS measurements: the measured chain reconfiguration time of ProTα in HeLa cells did not indicate folding or binding to an intracellular ligand or aggregation.

Combining single-molecule FRET spectroscopy with microinjection to deliver labeled proteins into living cell proved to be very versatile. We studied the intracellular folding kinetics of a destabilized variant of GB1 and the temperature dependent structural stability of frataxin (13). The folding kinetics of GB1 were analyzed using recurrence analysis (RASP), which enables the study of folding kinetics in a range of ~50 μs to ~100 ms (14). The difficulty of this experiment was to inject even lower concentrations of labeled protein than for regular single-molecule FRET experiments (target concentration 10 – 20 pM) and that data had to be acquired for many hours. However, owing to the confinement of molecules within the small cellular volume, and the resulting increased probability of recurrence compared to free solution, RASP turned out to be particularly well suited for in-cell data analysis. The result of this experiment showed that the intracellular folding kinetics in the millisecond range were not significantly altered compared to *in vitro* experiments. As a third example, the marginally stable protein frataxin was injected into HeLa cells to study the effect of the intracellular environment on its conformational stability. Frataxin shows both cold and heat denaturation in the range of 280 – 310 K, which makes it a good candidate to study heat denaturation in *in vivo* experiments. Indeed, it was possible to measure in-cell FRET histograms of frataxin from 279 – 299 K. Again, the results suggested that the conformational distributions and stability were not significantly altered by the cellular environment.

Although we studied many aspects of protein structures and dynamics in living cells with high sensitivity, the findings were always similar to results observed with the corresponding *in vitro* experiments, suggesting a surprising robustness of the conformational stabilities and folding kinetics of these proteins. The expectation was, however, that the complex mixture of macromolecules, metabolites and ions in cells would have a stronger effect on the proteins studied. An explanation for this behavior could be that the intracellular concentrations of macromolecules in eukaryotic cells is lower than the 300 to 400 mg/ml previously reported for E. coli (3), the value frequently assumed also for other cells. Our measurements indicated total protein concentrations of ~100 mg/ml and cytosolic protein concentrations of ~25 mg/ml (chapter 5), which is in agreement with other publications (4, 15, 16). Assuming cytosolic protein concentrations in this range, a comparison with our results from crowding experiments (2) indicate that the expected crowding effect on the dimensions of IDPs should in fact be very small. Increasing the intracellular concentrations of macromolecules might thus be a way to further elucidate the question of how crowded the cytosol is and whether intracellular conditions can be reached that alter IDP dimensions detectably.

We thus increased cellular crowding by subjecting HeLa cells to high concentrations of uncharged osmolytes in the medium (chapter 5). The osmotic shock should shrink the cell volume (17) and thus increase the intracellular concentrations of crowders. ProTα was used as a crowding sensor because of its known pronounced response to crowders *in vitro*. We indeed found that ProTα was slightly more compact in crowded cells compared to uncrowded cells. However, the chain reconfiguration time of the IDP was not affected (~80 ns in both crowded and uncrowded cells), suggesting that also under highly crowded conditions, ProTα remains unstructured and very flexible. On the other hand, translational diffusion of ProTα was ~4.5-fold slower compared to uncrowded conditions. By repeating translational diffusion measurements in uncrowded and crowded cells with other molecules of different sizes, we found that translational diffusion and chain reconfiguration were strongly length-scale dependent. Small molecules or the completely unstructured ProTα were much less affected by the crowded cytosol compared to rigid proteins in spite of a similar hydrodynamic radius as ProTα. We compared the in-cell results with *in vitro* experiments with differently sized PEGs at different concentrations. Assuming an average protein of the human proteome with a length of 375 amino acids (18) and globular shape, the hydrodynamic radius of ~3.5 nm (19) is in a similar regime as for PEG 6000 – 35’000. We thus used PEG 35’000 as a model crowder with concentrations of up to 20 % (w/v), which allowed us to mimic the crowded in-cell conditions well in terms of their effect on ProTα dimensions, dynamics and diffusion. The results are in accord with our measurements indicating that the average concentration of macromolecular crowders in the cytosol of unperturbed HeLa cells is rather low (~ 5 % w/v).

Altogether, we showed that confocal single-molecule FRET spectroscopy is a good method to study structural and dynamic properties of IDPs and other proteins in living cells. The single-molecule resolution of fluorescence spectroscopy in combination with microinjection for the delivery of labeled proteins into living cells has been found to be a very robust, reliable and versatile method to study proteins. We established very sensitive in-cell methods to study dimensions and dynamic processes of IDPs and other proteins in the sub-nanometer regime and as fast as nanoseconds, respectively. The high survival rate of injected cells and the intracellular stability of the fluorescent dyes should even allow us to do time-dependent in-cell experiments over several hours.

However, the methodology still offers room for improvement. First, a limitation of using a confocal microscope the way we employed it is that only freely diffusion proteins are studied. Unlike in free diffusion experiments, immobile proteins, e.g. bound to a membrane, will not get replenished after photobleaching. This effect reduces the observation time of an immobile species to a few milliseconds and thus makes it difficult to acquire enough data for a significant result. First attempts have been made to use scanning microscopy with a reduced laser intensity to analyze immobilized proteins in living cells (12). However, interpreting these fluorescence scans requires that the contribution of acceptor-inactive molecules and the fluorescence background are compensated appropriately.

Second, all our in-cell experiments were performed in HeLa cells, a cell line derived from cervix cancer tissue in 1951 (20). The use of HeLa cells as a model of human biology is questionable since its genome became highly randomized and mutated over the decades (21). Working with non-immortalized and healthy cells would thus be advantageous because it better represents biologically relevant systems. Neuronal cells derived from pluripotent stem cells (22), e.g., would be ideal to study the function of proteins such as amyloid-β or α-synuclein, which are related to Alzheimer's and Parkinson's disease, respectively (23).

A third issue is spatial resolution. The resolution of standard confocal detection is limited by diffraction to a few hundred nanometers, which currently allows us to distinguish between cytosolic or nuclear localization, but not smaller compartments. Increasing the spatial resolution, e.g. by making use of the recent developments in optical super-resolution microscopy, might enable better localization of proteins in organelles such as mitochondria or endosomes. Stimulated emission depletion (STED) microscopy, e.g., was combined with FCS to analyze membrane lipid dynamics in the plasma membrane of living cells (24) or membrane reorganization induced by pore-forming proteins (25) with a spatial resolution of less than 100 nm. The combination of STED with FRET is, however, technically difficult because the depletion laser used in STED will efficiently excite and thus bleach the acceptor dye of the FRET pair. Other super-resolution methods such as stochastic optical reconstruction microscopy (STORM) or photoactivation localization microscopy (PALM) provide rather low temporal resolution, so that dynamic processes faster than ~0.1 – 0.5 s currently cannot be studied (26). Also, a viable strategy to photoactivate donor and acceptor dyes at the same time has not been reported.

This limitations of spatial resolution for optical detection could be overcome in a combinatorial approach with other biophysical methods. Cryo-electron tomography (cryo-ET), e.g., can analyze macromolecular complexes in their native cellular setting with nanometer resolution. Detailed tomograms and structures of large protein complexes such as 26S proteasomes (27) or the nuclear pore complex (28, 29) were recorded in cells, even highlighting different functional states. Disadvantages of cryo-ET are that structural details of dynamic or disordered regions in protein complexes are averaged out during data analysis or get lost because samples are frozen or vitrified.

The resolution of FRET at the molecular scale can be further improved by combining it with nuclear magnetic resonance (NMR) spectroscopy, where sub-nanometer to atomic resolution can be obtained. Recent developments have made it possible to study isotope-labeled proteins also in living mammalian cells (30). NMR can resolve structural aspects over an ensemble of proteins inside cells; reports about long-range interactions with its surrounding; and provides dynamic information of backbone and side-chain motions. Especially small and soluble proteins can be characterized well by in-cell NMR. Using electroporation for protein delivery, the disordered character of α-synuclein (31) or the stability of the protein SOD1 (32) in mammalian cells was analyzed with NMR methods. Larger proteins or rigid protein complexes were successfully studied with in-cell solid-state NMR, which was, e.g., used to observe the structure and dynamics of epidermal growth factor receptor (EGFR) activation upon EGF stimulation in native membrane vesicles (33). Solid-state NMR can be especially useful for the analysis of larger protein oligomers or amyloid fibrils in cells (34).

A method similar to NMR is electron paramagnetic resonance (EPR) spectroscopy, where spin labels are introduced into the protein of interest via site-directed reactions. With two spin labels at different positions, distances of up to ~8 nm can be probed (35), but unlike single-molecule FRET spectroscopy, EPR results are ensemble-averaged over many cells and have to be acquired at cryogenic temperature for precise distance measurements. However, EPR has been used successfully to probe the architecture of α-synuclein fibrils (36) *in vitro* and the conformations of ubiquitin (37) in HeLa cells.

Our approach for analyzing intracellular proteins at the single-molecule level, in combination with biophysical methods such as NMR or cryo-ET, opens the field for further studies on protein structure, dynamics and function in living cells. The use of single-molecule FRET spectroscopy, SAXS and NMR *in vitro* has already been demonstrated in an integrated approach, e.g., by Aznauryan et al. (38) to study conformational and dynamic aspects of an unfolded protein. Milles et al. (39) used a combination of single-molecule FRET spectroscopy, NMR and molecular simulations to study the interaction between disordered nucleoporins in the nuclear pore complex with nuclear transport receptors. Combining single-molecule FRET with other in-cell methods will thus allow us to obtain insights into intracellular processes of folded and disordered proteins on a much more detailed level than possible to date.

References

1. Wright PE, Dyson HJ (2015) Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* 16(1):18–29.
2. Soranno A, et al. (2014) Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments. *Proc Natl Acad Sci USA* 111(13):4874–4879.
3. Zimmerman SB, Trach SO (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. *J Mol Biol* 222(3):599–620.
4. Cheung MC, et al. (2013) Intracellular protein and nucleic acid measured in eight cell types using deep-ultraviolet mass mapping. *Cytom Part A* 83(6):540–551.
5. Spitzer J, Poolman B (2013) How crowded is the prokaryotic cytoplasm? *FEBS Lett* 587(14):2094–2098.
6. Zhu L, Brangwynne CP (2015) Nuclear bodies: the emerging biophysics of nucleoplasmic phases. *Curr Opin Cell Biol* 34:23–30.
7. Dhar A, et al. (2011) Protein stability and folding kinetics in the nucleus and endoplasmic reticulum of eucaryotic cells. *Biophys J* 101(2):421–430.
8. Boersma AJ, Zuhorn IS, Poolman B (2015) A sensor for quantification of macromolecular crowding in living cells. *Nat Methods* 12(3):227–229.
9. Murakoshi H, et al. (2004) Single-molecule imaging analysis of Ras activation in living cells. *Proc Natl Acad Sci USA* 101(19):7317–7322.
10. Sakon JJ, Weninger KR (2010) Detecting the conformation of individual proteins in live cells. *Nat Methods* 7(3):203–205.
11. Crawford R, et al. (2013) Long-lived intracellular single-molecule fluorescence using electroporated molecules. *Biophys J* 105(11):2439–2450.
12. König I, et al. (2015) Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells. *Nat Methods* 12(8):773–779.
13. Aznauryan M, Nettels D, Holla A, Hofmann H, Schuler B (2013) Single-molecule spectroscopy of cold denaturation and the temperature-induced collapse of unfolded proteins. *J Am Chem Soc* 135(38):14040–14043.
14. Hoffmann A, et al. (2011) Quantifying heterogeneity and conformational dynamics from single molecule FRET of diffusing molecules: recurrence analysis of single particles (RASP). *Phys Chem Chem Phys* 13(5):1857–1871.
15. Phillip Y, Kiss V, Schreiber G (2012) Protein-binding dynamics imaged in a living cell. *Proc Natl Acad Sci USA* 109(5):1461–1466.
16. Wiśniewski JR, Hein MY, Cox J, Mann M (2014) A “proteomic ruler” for protein copy number and concentration estimation without spike-in standards. *Mol Cell Proteomics* 13(12):3497–3506.
17. Hoffmann EK, Lambert IH, Pedersen SF (2009) Physiology of Cell Volume Regulation in Vertebrates. *Physiol Rev* 89(1):193–277.
18. Brocchieri L, Karlin S (2005) Protein length in eukaryotic and prokaryotic

proteomes. *Nucleic Acids Res* 33(10):3390–3400.

19. Dill KA, Ghosh K, Schmit JD (2011) Physical limits of cells and proteomes. *Proc Natl Acad Sci USA* 108(44):17876–17882.
20. Scherer WF, Syverton JT, Gey GO (1953) Studies on the Propagation Invitro of Poliomyelitis Viruses .4. Viral Multiplication in a Stable Strain of Human Malignant Epithelial Cells (Strain Hela) Derived From an Epidermoid Carcinoma of the Cervix. *J Exp Med* 97(5):695–&.
21. Landry JJM, et al. (2013) The genomic and transcriptomic landscape of a HeLa cell line. *G3 (Bethesda)* 3(8):1213–1224.
22. Christian KM, Song H, Ming G-L (2014) Functions and dysfunctions of adult hippocampal neurogenesis. *Annu Rev Neurosci* 37:243–262.
23. Larson ME, et al. (2012) Soluble alpha-Synuclein Is a Novel Modulator of Alzheimer's Disease Pathophysiology. *J Neurosci* 32(30):10253–10266.
24. Eggeling C, et al. (2008) Direct observation of the nanoscale dynamics of membrane lipids in a living cell. *Nature* 457(7233):1159–1162.
25. Sarangi NK, P II, Ayappa KG, Visweswariah SS, Basu JK (2016) Super-resolution Stimulated Emission Depletion-Fluorescence Correlation Spectroscopy Reveals Nanoscale Membrane Reorganization Induced by Pore-Forming Proteins. *Langmuir* 32(37):9649–9657.
26. Jones SA, Shim S-H, He J, Zhuang X (2011) Fast, three-dimensional super-resolution imaging of live cells. *Nat Methods* 8(6):499–508.
27. Asano S, et al. (2015) Proteasomes. A molecular census of 26S proteasomes in intact neurons. *Science* 347(6220):439–442.
28. Eibauer M, et al. (2015) Structure and gating of the nuclear pore complex. *Nat Comms* 6:7532.
29. Appen von A, et al. (2015) In situ structural analysis of the human nuclear pore complex. *Nature* 526(7571):140–143.
30. Luchinat E, Banci L (2016) A Unique Tool for Cellular Structural Biology: In-cell NMR. *J Biol Chem* 291(8):3776–3784.
31. Theillet F-X, et al. (2016) Structural disorder of monomeric alpha-synuclein persists in mammalian cells. *Nature* 530:45–50.
32. Danielsson J, et al. (2015) Thermodynamics of protein destabilization in live cells. *Proc Natl Acad Sci USA* 112(40):12402–12407.
33. Kaplan M, et al. (2016) EGFR Dynamics Change during Activation in Native Membranes as Revealed by NMR. *Cell* 167(5):1241–1251.e11.
34. Qiang W, Yau W-M, Lu J-X, Collinge J, Tycko R (2017) Structural variation in amyloid-β fibrils from Alzheimer's disease clinical subtypes. *Nature* 541(7636):217–221.
35. Goldfarb D (2014) Gd3+ spin labeling for distance measurements by pulse EPR spectroscopy. *Phys Chem Chem Phys* 16(21):9685–9699.
36. Chen M, Margittai M, Chen J, Langen R (2007) Investigation of alpha-synuclein fibril structure by site-directed spin labeling. *J Biol Chem* 282(34):24970–24979.

37. Martorana A, et al. (2014) Probing Protein Conformation in Cells by EPR Distance Measurements using Gd3+ Spin Labeling. *J Am Chem Soc* 136(38):13458–13465.

38. Aznauryan M, et al. (2016) Comprehensive structural and dynamical view of an unfolded protein from the combination of single-molecule FRET, NMR, and SAXS. *Proc Natl Acad Sci USA* 113(37):E5389–98.

39. Milles S, et al. (2015) Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors. *Cell* 163(3):734–745.

Curriculum vitae

Personal Information

Name: Iwo König
Date of Birth: 24.02.1982
Nationality: Swiss

Education

02/2012 – present	University of Zurich **PhD studies** PhD thesis under the supervision of Prof. Dr. B. Schuler *Single-molecule FRET spectroscopy of intrinsically disordered proteins in crowded solutions and in living mammalian cells*
02/2010 – 02/2012	University of Zurich **MSc in Biochemistry** Master's thesis under the supervision of Prof. Dr. B. Schuler *Effect of crowding on the unfolded state dimensions of prothymosin α*
10/2002 – 12/2005	Zurich University of Applied Sciences **Chemist FH** Diploma thesis under the supervision of Prof. Dr. C. Zaborosch *Alternative purification process for monoclonal antibodies*
08/1998 – 08/2001	ETH Zurich **Chemistry lab technician**
08/1998 – 08/2001	Berufsmaturitätsschule Winterthur

Employment History

01/2006 – 12/2010	Zurich University of Applied Sciences **Scientific Assistant** Institute of Chemistry and Biological Chemistry, Prof. Dr. C. Zaborosch

Acknowledgements

I want to thank Prof. Ben Schuler for giving me the chance to do a PhD in his group. He allowed me to discover single-molecule spectroscopy and to develop new and exciting methods to study proteins in living cells. I enjoyed working under his supervision because he supported me with his expertise and he gave me the freedom to pursue projects according to my own ideas.

I also want to thank Prof. Andreas Plückthun and Prof. Achillefs Kapanidis for joining my thesis committee.

I especially want to thank Dr. Andrea Soranno for his supervision since my time as a Master’s student. During many discussions, he tried to give me insights into polymer physics, but more importantly, we supported me with uncountable ideas towards my project.

I thank Dr. Daniel Nettels for the time he invested to support my project with many lines of code and for teaching me how to use the microscopes.

My collaborators Dr. Arash Zarrine-Afsar and Dr. Mikayel Aznauryan deserve special thanks for their valuable contributions to my project and for becoming good friends.

My gratitude also goes to Jakob Stüber for teaching me how to work in a cell culture laboratory and for the help to get fluorescence images of cells. Furthermore, I want to thank the workshop team for building all the little tools I needed for my project.

I also want to thank all the other members of the Schuler group for the good working atmosphere and for the time we spent apart from work. I also want to thank my project students for their contributions and my Master’s student Raphael De Moliner for carrying out his work independently during the time I was writing my thesis.

Special thanks go to Prof. Christiane Zaborosch and to my former workgroup at ZHAW who paved the way for this career.

I finally want to thank my family for the enormous support and patience during my time as a PhD student.